GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:31:30 ; Search time 208.502 Seconds
(without alignments)
2194.362 Million cell updates/sec

Title: US-10-071-411A-1_COPY_950_1050
Perfect score: 101
Sequence: 1 ggaacctgggggtcaggccc.....cccgccgcgattcagagagc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883171

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	55	54.5	13249	15	US-10-311-455-90	Sequence 90, Appl
	2	43.2	42.8	13249	15	US-10-311-455-89	Sequence 89, Appl
	3	28.8	28.5	1989	15	US-10-156-761-7485	Sequence 7485, Ap
	4	28.8	28.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
c	5	28.6	28.3	1596	15	US-10-156-761-4833	Sequence 4833, Ap
c	6	28.6	28.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
	7	28.4	28.1	381	9	US-09-925-297-188	Sequence 188, App
c	8	28.2	27.9	884	9	US-09-790-045-11	Sequence 11, Appl
c	9	28.2	27.9	884	15	US-10-222-577-11	Sequence 11, Appl
c	10	28.2	27.9	884	15	US-10-222-578-11	Sequence 11, Appl
	11	27.8	27.5	412	9	US-09-833-381-649	Sequence 649, App
c	12	27.8	27.5	1735	16	US-10-159-563-190	Sequence 190, App
	13	27.8	27.5	2323	16	US-10-108-260A-2344	Sequence 2344, Ap
	14	27.8	27.5	2483	14	US-10-143-002-3	Sequence 3, Appli
	15	27.8	27.5	2483	15	US-10-325-891-3	Sequence 3, Appli
	16	27.8	27.5	14918	9	US-09-764-864-1766	Sequence 1766, Ap
	17	27.8	27.5	49152	13	US-10-087-192-646	Sequence 646, App
	18	27.6	27.3	87464	13	US-10-087-192-1411	Sequence 1411, Ap
c	19	27.4	27.1	366	9	US-09-796-692-5868	Sequence 5868, Ap
c	20	27.4	27.1	366	15	US-10-040-862-5868	Sequence 5868, Ap
c	21	27.4	27.1	366	16	US-10-057-475B-5868	Sequence 5868, Ap
c	22	27.4	27.1	366	16	US-10-154-884B-5868	Sequence 5868, Ap
	23	27.2	26.9	326	9	US-09-764-853-341	Sequence 341, App
	24	27.2	26.9	326	10	US-09-764-881-39	Sequence 39, Appl
	25	27.2	26.9	326	13	US-09-764-881-39	Sequence 39, Appl
	26	27.2	26.9	326	16	US-10-242-747-39	Sequence 39, Appl
c	27	27.2	26.9	669	15	US-10-156-761-7385	Sequence 7385, Ap
	28	27.2	26.9	993	13	US-10-282-122A-23817	Sequence 23817, A
c	29	27.2	26.9	1191	15	US-10-156-761-1010	Sequence 1010, Ap
	30	27.2	26.9	144179	13	US-10-087-192-511	Sequence 511, App
c	31	27	26.7	620	13	US-10-027-632-65038	Sequence 65038, A
c	32	27	26.7	620	16	US-10-027-632-65038	Sequence 65038, A
	33	27	26.7	846	15	US-10-156-761-7276	Sequence 7276, Ap
c	34	27	26.7	2003	10	US-09-933-767-171	Sequence 171, App
c	35	27	26.7	2003	13	US-10-004-860-171	Sequence 171, App
c	36	27	26.7	2003	15	US-10-023-282-171	Sequence 171, App
c	37	27	26.7	2419	16	US-10-108-260A-1507	Sequence 1507, Ap
	38	27	26.7	4608	13	US-10-342-887-14	Sequence 14, Appl
	39	27	26.7	4608	13	US-10-172-118-14	Sequence 14, Appl
	40	27	26.7	13460	9	US-09-764-868-1358	Sequence 1358, Ap
c	41	27	26.7	17967	9	US-09-764-848-50	Sequence 50, Appl
c	42	27	26.7	17967	15	US-10-116-016-50	Sequence 50, Appl
	43	27	26.7	17967	15	US-10-103-313-650	Sequence 650, App
c	44	27	26.7	17967	15	US-10-222-020-50	Sequence 50, Appl
c	45	27	26.7	32659	16	US-10-085-117-208	Sequence 208, App

ALIGNMENTS

RESULT 1
US-10-311-455-90/c
; Sequence 90, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 90
; LENGTH: 13249
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-90

Query Match 54.5%; Score 55; DB 15; Length 13249;
Best Local Similarity 78.8%; Pred. No. 1.3e-08;
Matches 78; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

```
Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||   |||  ||||| |||||  ||  | ||||||  | ||||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
```

```
                ||||||||||| |||||| |||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992
```

RESULT 2
US-10-311-455-89
; Sequence 89, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 89
; LENGTH: 13249
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-89

Query Match 42.8%; Score 43.2; DB 15; Length 13249;
Best Local Similarity 71.0%; Pred. No. 8.9e-05;
Matches 71; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

```
Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||  ||||||| |||    || ||||||||| ||   |||||||||||||| || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  ||| |||||  || |||||| ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257
```

RESULT 3
US-10-156-761-7485
; Sequence 7485, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7485
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1989)

US-10-156-761-7485

Query Match 28.5%; Score 28.8; DB 15; Length 1989;
Best Local Similarity 58.0%; Pred. No. 4.8;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy          2 GAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTC 61
              || || || | ||   ||| |   |||| ||  | || ||  ||||||| ||  ||
Db        444 GACCCCGGCGATCGTCCCCATGGGCCGGGCGGCCGGACCGGGCTCGCGCGACGCCCGCTT 503

Qy         62 CACACCCTTCCAGGCATTTGCCCGCCGC 89
              ||  || | | ||||   || ||   ||
Db        504 CATCCCGTACGAGGCCCATGACCCGGGC 531
```

RESULT 4
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 28.5%; Score 28.8; DB 15; Length 9025608;
Best Local Similarity 58.0%; Pred. No. 2.8;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy          2 GAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTC 61
              || || || | ||   ||| |   |||| ||  | || ||  ||||||| ||  ||
Db    8949160 GACCCCGGCGATCGTCCCCATGGGCCGGGCGGCCGGACCGGGCTCGCGCGACGCCCGCTT 8949219

Qy         62 CACACCCTTCCAGGCATTTGCCCGCCGC 89
              ||  || | | ||||   || ||   ||
Db    8949220 CATCCCGTACGAGGCCCATGACCCGGGC 8949247
```

RESULT 5
US-10-156-761-4833/c
; Sequence 4833, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262

```
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4833
;   LENGTH: 1596
;   TYPE: DNA
;   ORGANISM: Streptomyces avermitilis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1596)
US-10-156-761-4833

  Query Match             28.3%;  Score 28.6;  DB 15;  Length 1596;
  Best Local Similarity   57.1%;  Pred. No. 5.7;
  Matches   52;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy          4 ACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCA 63
              ||   || || |||||||| |    || | | ||   |  |  |  |||  |||||||| ||
Db       1252 ACAGGGTGGCCAGGCCCCCGATCAGGAACGTGCCGGCGAGCACCGGCGCGACCTTCTGCA 1193

Qy         64 CACCCTTCCAGGCATTTGCCCGCCGCGATTC 94
              | |||    || |        |||||||| |
Db       1192 CGCCCCCGTAGTCGGCGATGAGCCGCGATCC 1162


RESULT 6
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
;   LENGTH: 9025608
;   TYPE: DNA
;   ORGANISM: Streptomyces avermitilis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4187715)
;   OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

  Query Match             28.3%;  Score 28.6;  DB 15;  Length 9025608;
  Best Local Similarity   57.1%;  Pred. No. 3.2;
  Matches   52;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy          4 ACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCA 63
              ||   || || |||||||| |    || | | ||   |  |  |  |||  |||||||| ||
Db    5894377 ACAGGGTGGCCAGGCCCCCGATCAGGAACGTGCCGGCGAGCACCGGCGCGACCTTCTGCA 5894318

Qy         64 CACCCTTCCAGGCATTTGCCCGCCGCGATTC 94
              | |||    || |        |||||||| |
Db    5894317 CGCCCCCGTAGTCGGCGATGAGCCGCGATCC 5894287


RESULT 7
US-09-925-297-188
; Sequence 188, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA105
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05989
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 188
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (293)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (350)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-188

  Query Match             28.1%;  Score 28.4;  DB 9;  Length 381;
  Best Local Similarity   53.7%;  Pred. No. 7.3;
  Matches   44;  Conservative    6;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         18 CCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTTCCAGGCA 77
              |:|||:  || || || || |||:|    |: |   | || |  |: || |  :  || ||
Db         96 CYCCARGTGCAGGCAGAGCCCCCRGAGMCRTGGCCAGCCCTTCCSGCAGCTCYGAAGCCA 155

Qy         78 TTTGCCCGCCGCGATTCAGAGA 99
                | ||  ||| |||  ||| ||
Db        156 CTGGCAAGCCCCGAGGCAGGGA 177


RESULT 8
US-09-790-045-11/c
; Sequence 11, Application US/09790045
; Patent No. US20020052047A1
; GENERAL INFORMATION:
;  APPLICANT: Hasebe, Akira
;  APPLICANT:  Tsuchiya, Kenichi
;  APPLICANT:  Horita, Mitsuo
;  TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia Solanacearum
;  FILE REFERENCE: NANP108US
;  CURRENT APPLICATION NUMBER: US/09/790,045
;  CURRENT FILING DATE:  2001-02-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 884
;   TYPE: DNA
;   ORGANISM: Ralstonia solanacearum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (44)...(865)
US-09-790-045-11

  Query Match             27.9%;  Score 28.2;  DB 9;  Length 884;
  Best Local Similarity   59.3%;  Pred. No. 8;
  Matches   48;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  ||||  ||
Db        255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196
```

```
Qy          69 TTCCAGGCATTTGCCCGCCGC 89
               || || |     |||| |||
Db         195 CTCAAGTCGCACTCCCGACGC 175
```

RESULT 9
US-10-222-577-11/c
; Sequence 11, Application US/10222577
; Publication No. US20030009026A1
; GENERAL INFORMATION:
; APPLICANT: Hasebe, Akira
; APPLICANT: Tsuchiya, Kenichi
; APPLICANT: Horita, Mitsuo
; TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia
; TITLE OF INVENTION: Solanacearum
; FILE REFERENCE: NANP108US
; CURRENT APPLICATION NUMBER: US/10/222,577
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/790,045
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Ralstonia solanacearum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)...(865)
US-10-222-577-11

Query Match 27.9%; Score 28.2; DB 15; Length 884;
Best Local Similarity 59.3%; Pred. No. 8;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy           9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
               ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  |||| ||
Db         255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196

Qy          69 TTCCAGGCATTTGCCCGCCGC 89
               || || |     |||| |||
Db         195 CTCAAGTCGCACTCCCGACGC 175
```

RESULT 10
US-10-222-578-11/c
; Sequence 11, Application US/10222578
; Publication No. US20030027340A1
; GENERAL INFORMATION:
; APPLICANT: Hasebe, Akira
; APPLICANT: Tsuchiya, Kenichi
; APPLICANT: Horita, Mitsuo
; TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia
; TITLE OF INVENTION: Solanacearum
; FILE REFERENCE: NANP108US
; CURRENT APPLICATION NUMBER: US/10/222,578
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/790,045
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Ralstonia solanacearum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)...(865)
US-10-222-578-11

Query Match 27.9%; Score 28.2; DB 15; Length 884;
Best Local Similarity 59.3%; Pred. No. 8;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy           9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
               ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  |||| ||
Db         255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196

Qy          69 TTCCAGGCATTTGCCCGCCGC 89
               || || |     |||| |||
Db         195 CTCAAGTCGCACTCCCGACGC 175
```

RESULT 11
US-09-833-381-649
; Sequence 649, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 649
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-649

Query Match 27.5%; Score 27.8; DB 9; Length 412;
Best Local Similarity 57.5%; Pred. No. 11;
Matches 50; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy          11 GGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTT 70
               || |||||||| | | |  || ||| |||| |||| | |   | |||| |||
Db         144 GGCCAGGCCCCCCCGGAGATGAGGGCGGACAGGCCCGCGAGGACGGGTGCCCACGCCCCA 203

Qy          71 CCAGGCATTTGCCCGCCGCGATTCAGA 97
               | |||   |  || |||   ||||
Db         204 CACGGCGAAGTCTAGCTGCGCCCCAGA 230
```

RESULT 12
US-10-159-563-190/c
; Sequence 190, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56USI1
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 190
; LENGTH: 1735
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-190

Query Match 27.5%; Score 27.8; DB 16; Length 1735;
Best Local Similarity 59.5%; Pred. No. 10;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy           9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
               || |   ||| || || ||| |  || | |       |||| |||  || | |||||  |
```

```
Db        117 GGCGCTCGGCGCCTGCTGCGTGCTGCCCTCTTCCCCGCGCACGATGCCGTGTCCACCGCT 58

Qy         69 TTCCAGGCATTTGCCCGCC 87
              ||||||| |  ||| ||
Db         57 GTCCAGGCTCTGTCCCTCC 39


RESULT 13
US-10-108-260A-2344
; Sequence 2344, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2344
;   LENGTH: 2323
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-108-260A-2344

  Query Match             27.5%;  Score 27.8;  DB 16;  Length 2323;
  Best Local Similarity   74.5%;  Pred. No. 10;
  Matches   35;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCG 51
              ||   ||||| |||||||||||| ||    |||||||||| | ||| |
Db        148 CCGCAGGGTCCGGCCCCAGCCGGGGTCCCCGCGCCCAGGCGAGCGGG 194


RESULT 14
US-10-143-002-3
; Sequence 3, Application US/10143002
; Publication No. US20020132775A1
;   GENERAL INFORMATION:
;        APPLICANT: LI, ET AL.
;        TITLE OF INVENTION: Potassium Channel Protein 1
;                            and 2
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
;                        CECCHI, STEWART & OLSTEIN
;             STREET: 6 BECKER FARM ROAD
;             CITY: ROSELAND
;             STATE: NEW JERSEY
;             COUNTRY: USA
;             ZIP: 07068
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 INCH DISKETTE
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: MS-DOS
;             SOFTWARE: WORD PERFECT 5.1
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/143,002
;             FILING DATE: 13-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/102,493
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US/08/468,533
;             FILING DATE: 6 JUNE 1995
;             APPLICATION NUMBER: PCT/US94/08449
;             FILING DATE: 28 JUL 1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: MULLINS, J.G.
;             REGISTRATION NUMBER: 33,073
;             REFERENCE/DOCKET NUMBER: 325800-310
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-994-1700
;             TELEFAX: 201-994-1744
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2483 BASE PAIRS
;             TYPE: NUCLEIC ACID
;             STRANDEDNESS: SINGLE
;             TOPOLOGY: LINEAR
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-143-002-3

  Query Match             27.5%;  Score 27.8;  DB 14;  Length 2483;
  Best Local Similarity   57.5%;  Pred. No. 10;
  Matches   50;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || |  | |   |||||| || |   ||| || || ||  ||    || |  | |
Db        455 CCGGGCGCCCGGATGCCAGCCCCGAGCCCCGCCGCCGGGTGCATGCCTCCCCCGCGGCGC 514

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGCGA 91
              |||   |||||  |||||||  | ||
Db        515 GCCCCCGCAGGCTGCTGCCCGCTGTGA 541


RESULT 15
US-10-325-891-3
; Sequence 3, Application US/10325891
; Publication No. US20030092895A1
;   GENERAL INFORMATION:
;        APPLICANT: LI, ET AL.
;        TITLE OF INVENTION: Human Potassium Channel 1 and 2 Proteins
;        NUMBER OF SEQUENCES: 13
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
;                        CECCHI, STEWART & OLSTEIN
;             STREET: 6 BECKER FARM ROAD
;             CITY: ROSELAND
;             STATE: NEW JERSEY
;             COUNTRY: USA
;             ZIP: 07068
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 INCH DISKETTE
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: MS-DOS
;             SOFTWARE: WORD PERFECT 5.1
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/325,891
;             FILING DATE: 23-Dec-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/009,492
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/464,340
;             FILING DATE: June 5,1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: FERRARO, GREGORY D.
;             REGISTRATION NUMBER: 36,134
;             REFERENCE/DOCKET NUMBER: 325800-415
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-994-1700
;             TELEFAX: 201-994-1744
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2483 BASE PAIRS
;             TYPE: NUCLEIC ACID
;             STRANDEDNESS: SINGLE
;             TOPOLOGY: LINEAR
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-325-891-3

  Query Match             27.5%;  Score 27.8;  DB 15;  Length 2483;
  Best Local Similarity   57.5%;  Pred. No. 10;
```

```
  Matches   50;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || |  | |   |||||| || |   |||  || || ||  ||    ||  |  | |
Db        455 CCGGGCGCCCGGATGCCAGCCCCGAGCCCCGCCGCCGGGTGCATGCCTCCCCCGCGGCGC 514

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGCGA 91
                |||   |||||   |||||||  | ||
Db        515 GCCCCCGCAGGCTGCTGCCCGCTGTGA 541


Search completed: May  7, 2004, 17:35:50
Job time : 454.502 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:35:03 ; Search time 50.7167 Seconds
(without alignments)
1105.159 Million cell updates/sec

Title: US-10-071-411A-1_COPY_950_1050
Perfect score: 101
Sequence: 1 ggaacctgggggtcaggccc.....cccgccgcgattcagagagc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	28.2	27.9	884	4	US-09-790-045-11	Sequence 11, Appl
c	2	28.2	27.9	884	4	US-10-222-577-11	Sequence 11, Appl
c	3	28.2	27.9	884	4	US-10-222-578-11	Sequence 11, Appl
	4	27.8	27.5	412	4	US-09-833-381-649	Sequence 649, App
	5	27.8	27.5	2483	1	US-08-464-340A-3	Sequence 3, Appli
	6	27.8	27.5	2483	5	PCT-US94-08449A-3	Sequence 3, Appli
	7	27.4	27.1	1665	4	US-09-489-039A-831	Sequence 831, App
c	8	27	26.7	2003	4	US-09-205-258-171	Sequence 171, App
	9	26.6	26.3	599	4	US-09-621-976-2262	Sequence 2262, Ap
c	10	26.6	26.3	1326	4	US-09-252-991A-15544	Sequence 15544, A
	11	26.6	26.3	1437	4	US-09-252-991A-15295	Sequence 15295, A
	12	26.4	26.1	591	4	US-09-252-991A-13956	Sequence 13956, A
c	13	26.4	26.1	742	3	US-09-085-848-2	Sequence 2, Appli
c	14	26.4	26.1	742	4	US-09-562-616-2	Sequence 2, Appli
c	15	26.4	26.1	2103	4	US-09-252-991A-13604	Sequence 13604, A
c	16	26.4	26.1	2259	4	US-09-252-991A-13396	Sequence 13396, A
	17	26.4	26.1	3222	4	US-09-252-991A-13746	Sequence 13746, A
c	18	26.2	25.9	807	4	US-09-489-039A-5230	Sequence 5230, Ap
	19	26	25.7	647	1	US-08-260-202A-9	Sequence 9, Appli
	20	26	25.7	647	1	US-08-017-114-9	Sequence 9, Appli
	21	26	25.7	647	3	US-08-505-307-9	Sequence 9, Appli
	22	26	25.7	647	4	US-09-609-151A-9	Sequence 9, Appli
	23	26	25.7	647	5	PCT-US94-02034-9	Sequence 9, Appli
c	24	26	25.7	828	4	US-09-252-991A-13980	Sequence 13980, A
	25	26	25.7	840	4	US-09-252-991A-13069	Sequence 13069, A
	26	26	25.7	1605	4	US-09-252-991A-14154	Sequence 14154, A
	27	26	25.7	1824	4	US-09-252-991A-14243	Sequence 14243, A
c	28	26	25.7	6803	3	US-08-665-259-19	Sequence 19, Appl
c	29	26	25.7	6803	3	US-08-762-500-19	Sequence 19, Appl
c	30	26	25.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	31	26	25.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	32	25.8	25.5	439	4	US-09-621-976-12886	Sequence 12886, A
c	33	25.8	25.5	876	4	US-09-252-991A-11208	Sequence 11208, A
	34	25.8	25.5	945	4	US-09-252-991A-11046	Sequence 11046, A
c	35	25.8	25.5	1098	4	US-09-252-991A-15756	Sequence 15756, A
	36	25.8	25.5	1356	4	US-09-252-991A-9233	Sequence 9233, Ap
	37	25.8	25.5	1923	4	US-09-252-991A-10971	Sequence 10971, A
c	38	25.8	25.5	2715	4	US-09-252-991A-8715	Sequence 8715, Ap
	39	25.8	25.5	3126	4	US-09-252-991A-9066	Sequence 9066, Ap
c	40	25.8	25.5	3226	3	US-08-870-126-10	Sequence 10, Appl
c	41	25.8	25.5	3226	4	US-09-445-247-10	Sequence 10, Appl
	42	25.8	25.5	3402	4	US-09-252-991A-15560	Sequence 15560, A
c	43	25.8	25.5	3606	4	US-09-252-991A-15688	Sequence 15688, A
	44	25.8	25.5	29629	4	US-09-729-995-3	Sequence 3, Appli
	45	25.8	25.5	29629	4	US-10-135-689-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-790-045-11/c
; Sequence 11, Application US/09790045
; Patent No. 6492510
; GENERAL INFORMATION:
; APPLICANT: Hasebe, Akira
; APPLICANT: Tsuchiya, Kenichi
; APPLICANT: Horita, Mitsuo
; TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia Solanacearum
; Patent No. 6492510
; FILE REFERENCE: NANP108US
; CURRENT APPLICATION NUMBER: US/09/790,045
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Ralstonia solanacearum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)...(865)
US-09-790-045-11

Query Match 27.9%; Score 28.2; DB 4; Length 884;
Best Local Similarity 59.3%; Pred. No. 7.2;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  |||| ||
Db        255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196

Qy         69 TTCCAGGCATTTGCCCGCCGC 89
              || || |     |||| |||
Db        195 CTCAAGTCGCACTCCCGACGC 175
```

RESULT 2
US-10-222-577-11/c
; Sequence 11, Application US/10222577
; Patent No. 6538125
; GENERAL INFORMATION:
; APPLICANT: Hasebe, Akira
; APPLICANT: Tsuchiya, Kenichi
; APPLICANT: Horita, Mitsuo
; TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia
; Patent No. 6538125
; TITLE OF INVENTION: Solanacearum
; FILE REFERENCE: NANP108US
; CURRENT APPLICATION NUMBER: US/10/222,577

```
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: US/09/790,045
;  PRIOR FILING DATE: 2001-02-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 884
;   TYPE: DNA
;   ORGANISM: Ralstonia solanacearum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (44)...(865)
US-10-222-577-11

  Query Match            27.9%;  Score 28.2;  DB 4;  Length 884;
  Best Local Similarity  59.3%;  Pred. No. 7.2;
  Matches   48;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  |||| ||
Db        255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196

Qy         69 TTCCAGGCATTTGCCCGCCGC 89
              || || |    |||| |||
Db        195 CTCAAGTCGCACTCCCGACGC 175


RESULT 3
US-10-222-578-11/c
; Sequence 11, Application US/10222578
; Patent No. 6570007
; GENERAL INFORMATION:
;  APPLICANT: Hasebe, Akira
;  APPLICANT:  Tsuchiya, Kenichi
;  APPLICANT:  Horita, Mitsuo
;  TITLE OF INVENTION: Insertion Sequence Element Derived From Ralstonia
; Patent No. 6570007
;  TITLE OF INVENTION:  Solanacearum
;  FILE REFERENCE: NANP108US
;  CURRENT APPLICATION NUMBER: US/10/222,578
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: US/09/790,045
;  PRIOR FILING DATE: 2001-02-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 884
;   TYPE: DNA
;   ORGANISM: Ralstonia solanacearum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (44)...(865)
US-10-222-578-11

  Query Match            27.9%;  Score 28.2;  DB 4;  Length 884;
  Best Local Similarity  59.3%;  Pred. No. 7.2;
  Matches   48;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              ||   || || ||| |||| |   || ||| |  ||  |||| |  | ||  |||| ||
Db        255 GGCAGCATGCGCCAACCGCAGCCCGCTCGCACCAGATAGCGCAAGGCGTTGACCACCTCC 196

Qy         69 TTCCAGGCATTTGCCCGCCGC 89
              || || |    |||| |||
Db        195 CTCAAGTCGCACTCCCGACGC 175


RESULT 4
US-09-833-381-649
; Sequence 649, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 649
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-649

  Query Match            27.5%;  Score 27.8;  DB 4;  Length 412;
  Best Local Similarity  57.5%;  Pred. No. 8.5;
  Matches   50;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         11 GGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTT 70
              || ||||||||  | | |   || |||  ||||  |||| | |    |  |||| |||
Db        144 GGCCAGGCCCCCCCCGGAGATGAGGGCGGACAGGCCCGCGAGGACGGGTGCCCACGCCCCA 203

Qy         71 CCAGGCATTTGCCCGCCGCGATTCAGA 97
                 |  |||      |  || |||   ||||
Db        204 CACGGCGAAGTCTAGCTGCGCCCCAGA 230


RESULT 5
US-08-464-340A-3
; Sequence 3, Application US/08464340A
; Patent No. 5710019
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Potassium Channel 1 and 2 Proteins
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,340A
;      FILING DATE:  June 5,1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/08449
;      FILING DATE:  28 JUL 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2483 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
US-08-464-340A-3

  Query Match            27.5%;  Score 27.8;  DB 1;  Length 2483;
  Best Local Similarity  57.5%;  Pred. No. 11;
```

```
  Matches   50;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || |  | |   ||||| || |   |||  || || ||  ||    ||  | | |
Db        455 CCGGGCGCCCGGATGCCAGCCCCGAGCCCCGCCGCCGGGTGCATGCCTCCCCCGCGGCGC 514

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGCGA 91
                |||   |||||   |||||||  | ||
Db        515 GCCCCCGCAGGCTGCTGCCCGCTGTGA 541


RESULT 6
PCT-US94-08449A-3
; Sequence 3, Application PC/TUS9408449A
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Potassium Channel Protein 1 and 2
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/08449A
;      FILING DATE:  SUBMITTED HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-105
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2483 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
PCT-US94-08449A-3

  Query Match             27.5%;  Score 27.8;  DB 5;  Length 2483;
  Best Local Similarity   57.5%;  Pred. No. 11;
  Matches   50;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || |  | |   ||||| || |   |||  || || ||  ||    ||  | | |
Db        455 CCGGGCGCCCGGATGCCAGCCCCGAGCCCCGCCGCCGGGTGCATGCCTCCCCCGCGGCGC 514

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGCGA 91
                |||   |||||   |||||||  | ||
Db        515 GCCCCCGCAGGCTGCTGCCCGCTGTGA 541


RESULT 7
US-09-489-039A-831
; Sequence 831, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 831
;   LENGTH: 1665
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-831

  Query Match             27.1%;  Score 27.4;  DB 4;  Length 1665;
  Best Local Similarity   54.5%;  Pred. No. 14;
  Matches   55;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||   | || |||||   ||    |  |  | || ||||| ||| || |  || | | ||
Db        762 GGCCGCCGGTGGTCACAACCTGCTCCTGATAGGCCCGCCCGGGACCGGGAAAACCATGCT 821

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
                 || ||  |     |  |  || ||||| ||| ||  | |
Db        822 AGCCAGCCGACTCCCCGGTCTCCTGCCGCCATTAAGCAATC 862


RESULT 8
US-09-205-258-171/c
; Sequence 171, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
```

```
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
;   LENGTH: 2003
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1961)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1999)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-171

  Query Match             26.7%;  Score 27;  DB 4;  Length 2003;
  Best Local Similarity   57.8%;  Pred. No. 19;
  Matches   48;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              || |||| | || ||| | || ||| ||| |||  | |||||  |  |  |||  |
Db         83 GGGACCTTGAGGGCAGCCACCGGCCCCGGACAGCAGCCTCAGGACAGGACCTCACCGATT 24

Qy         61 CCACACCCTTCCAGGCATTTGCC 83
              || |   |  | |||  ||||
Db         23 CCTCCTGCAAGCTGGCTCGTGCC 1
```

```
RESULT 9
US-09-621-976-2262
; Sequence 2262, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 2262
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 216..368
US-09-621-976-2262

  Query Match             26.3%;  Score 26.6;  DB 4;  Length 599;
  Best Local Similarity   54.6%;  Pred. No. 21;
  Matches   53;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              | ||  |||  || ||||| |||          | |   || | ||||    ||  |||| |
Db        257 CATGCTGGTTCGGTCCCAGACGCTCAGGTGCCACATTGGGGTGCGCTCTCCCGGCTCCTC 316

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
                |   | | |||  ||||||| | |    ||| |||
Db        317 CACACGCTACGCAGCTGCCCGCGGGGTGAAAGAAAGC 353
```

```
RESULT 10
US-09-252-991A-15544/c
; Sequence 15544, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 15544
;   LENGTH: 1326
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (92)
;   OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15544

  Query Match             26.3%;  Score 26.6;  DB 4;  Length 1326;
  Best Local Similarity   58.0%;  Pred. No. 23;
  Matches   47;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || |  |  | ||  ||  ||| ||  |||||  |||  | || ||| || || | | ||
Db       1215 CCGGTCGACCTGGAGCCGACCGTGGTCAGCGCCACCACTACCGAGCGCAAGCTGCGCGAC 1156

Qy         65 ACCCTTCCAGGCATTTGCCCG 85
                |||   |  || |||  | |
Db       1155 GCCCCGGCCAGCGTTTCGCTG 1135
```

```
RESULT 11
US-09-252-991A-15295
; Sequence 15295, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 15295
;   LENGTH: 1437
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1253)
;   OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-15295

  Query Match             26.3%;  Score 26.6;  DB 4;  Length 1437;
  Best Local Similarity   58.0%;  Pred. No. 24;
  Matches   47;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || |  |  | ||  ||  ||| ||  ||||  |||  | || ||| || || | | ||
Db        130 CCGGTCGACCTGGAGCCGACCGTGGTCAGCGCCACCACTACCGAGCGCAAGCTGCGCGAC 189

Qy         65 ACCCTTCCAGGCATTTGCCCG 85
                |||   |  || |||  | |
Db        190 GCCCCGGCCAGCGTTTCGCTG 210


RESULT 12
US-09-252-991A-13956
; Sequence 13956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13956
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13956

  Query Match             26.1%;  Score 26.4;  DB 4;  Length 591;
  Best Local Similarity   55.4%;  Pred. No. 24;
  Matches   51;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          8 GGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACC 67
              ||  | | || ||   ||   ||  |||||| || || ||||| |  |     |  || |
Db         81 GGCAGGCTGGGCCAGTCCAGCGGCTGCGCGCGCAAGACCGCGCCAGGCAGCAGCGCCAGC 140

Qy         68 CTTCCAGGCATTTGCCCGCCGCGATTCAGAGA 99
                |||    ||  ||  || |||  |||| ||
Db        141 AGTCCGCCGATCAGCAGGCGGCGCATCAGGGA 172


RESULT 13
US-09-085-848-2/c
; Sequence 2, Application US/09085848
; Patent No. 6171855
;  GENERAL INFORMATION:
;    APPLICANT:  ASKARI, FREDERICK K.
;    TITLE OF INVENTION:  IMPROVED VECTORS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/085,848
;      FILING DATE:  28-MAY-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARROLL, PETER G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  UM-03338
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  742 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
US-09-085-848-2

  Query Match             26.1%;  Score 26.4;  DB 3;  Length 742;
  Best Local Similarity   59.2%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         12 GTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTTC 71
              |||    |||||| |||||   | | | |||||    | |  | |  |   || ||| ||
Db        496 GTCCATGCCCAGCAGCGGGGCCCACCCACAGGACGTGGGGGTATCCCCGGAACCCCCATC 437

Qy         72 CAGGCATTTGCCCGCC 87
              | |  || ||| || ||
Db        436 CGGACACTTGTCCCCC 421


RESULT 14
US-09-562-616-2/c
; Sequence 2, Application US/09562616
; Patent No. 6599744
;   GENERAL INFORMATION:
;        APPLICANT: ASKARI, FREDERICK K.
;        TITLE OF INVENTION: IMPROVED VECTORS
;        NUMBER OF SEQUENCES: 2
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MEDLEN & CARROLL, LLP
;             STREET: 220 MONTGOMERY STREET, SUITE 2200
;             CITY: SAN FRANCISCO
;             STATE: CALIFORNIA
;             COUNTRY: UNITED STATES OF AMERICA
;             ZIP: 94104
;        COMPUTER READABLE FORM:
```

```
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/562,616
;             FILING DATE: 01-May-2000
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/085,848
;             FILING DATE: 28-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;             NAME: CARROLL, PETER G.
;             REGISTRATION NUMBER: 32,837
;             REFERENCE/DOCKET NUMBER: UM-03338
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 742 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-562-616-2

  Query Match             26.1%;  Score 26.4;  DB 4;  Length 742;
  Best Local Similarity   59.2%;  Pred. No. 25;
  Matches   45;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         12 GTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTTC 71
              |||    |||||| |||||   | | | |||||    | |  | |  |   || ||| ||
Db        496 GTCCATGCCCAGCAGCGGGGCCCACCCACAGGACGTGGGGGTATCCCCGGAACCCCCATC 437

Qy         72 CAGGCATTTGCCCGCC 87
              | | || ||| || ||
Db        436 CGGACACTTGTCCCCC 421


RESULT 15
US-09-252-991A-13604/c
; Sequence 13604, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13604
;   LENGTH: 2103
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13604

  Query Match             26.1%;  Score 26.4;  DB 4;  Length 2103;
  Best Local Similarity   55.4%;  Pred. No. 29;
  Matches   51;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          8 GGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACC 67
              ||  | | ||  ||   ||   ||  |||||| || || ||||| |  |     |  || |
Db       1242 GGCAGGCTGGGCCAGTCCAGCGGCTGCGCGCGCAAGACCGCGCCAGGCAGCAGCGCCAGC 1183

Qy         68 CTTCCAGGCATTTGCCCGCCGCGATTCAGAGA 99
                 |||    ||  ||  || |||  |||| ||
Db       1182 AGTCCGCCGATCAGCAGGCGGCGCATCAGGGA 1151


Search completed: May  7, 2004, 15:44:40
Job time : 57.7167 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 11:56:28 ; Search time 206.335 Seconds
(without alignments)
2079.475 Million cell updates/sec

Title: US-10-071-411A-1_COPY_950_1050
Perfect score: 101
Sequence: 1 ggaacctgggggtcaggccc.....cccgccgcgattcagagagc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747720

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	89	88.1	168174	6	ABT11173	Abt11173 Human 5-l
	2	89	88.1	168273	6	ABT11114	Abt11114 Human 5-l
c	3	55	54.5	13249	6	ABL32117	Abl32117 Human imm
c	4	55	54.5	13249	6	ABK31177	Abk31177 Signal tr
c	5	55	54.5	13249	6	ABL70132	Abl70132 Chemicall
	6	43.2	42.8	13249	6	ABL32116	Abl32116 Human imm
	7	43.2	42.8	13249	6	ABK31176	Abk31176 Signal tr
	8	43.2	42.8	13249	6	ABL70131	Abl70131 Chemicall
c	9	30	29.7	44147	6	ABK84481	Abk84481 Human cDN
c	10	30	29.7	44147	9	ADD14691	Add14691 Human src
c	11	29.4	29.1	33737	6	AAL50815	Aal50815 Human can
	12	29	28.7	956	7	ABZ59005	Abz59005 Human onc
c	13	29	28.7	990	7	ABZ52368	Abz52368 Aspergill
	14	29	28.7	3761	7	ABZ58996	Abz58996 Human onc
	15	29	28.7	10735	7	ABZ58995	Abz58995 Human onc
	16	28.4	28.1	381	3	AAC98960	Aac98960 Human pan
c	17	28.2	27.9	890	6	ABK15293	Abk15293 Ralstonia
c	18	28	27.7	38064	6	ABA99469	Aba99469 Actinopla
c	19	27.8	27.5	1735	2	AAT77840	Aat77840 Human mel
	20	27.8	27.5	2483	2	AAT12462	Aat12462 Human K+
	21	27.8	27.5	2483	2	AAV04874	Aav04874 cDNA sequ
	22	27.8	27.5	2483	9	ADC99154	Adc99154 Human mat
c	23	27.8	27.5	4456	7	ADA57655	Ada57655 BAC fragm
c	24	27.8	27.5	4456	7	ADA41527	Ada41527 Human sec
c	25	27.8	27.5	4456	9	ADC74634	Adc74634 Human sec
	26	27.8	27.5	14918	4	AAS26792	Aas26792 Human gen
c	27	27.8	27.5	14918	7	ADA57654	Ada57654 BAC fragm
c	28	27.8	27.5	14918	7	ADA41526	Ada41526 Human sec
	29	27.8	27.5	14918	7	ABX74141	Abx74141 Human nov
c	30	27.8	27.5	14918	9	ADC74633	Adc74633 Human sec
	31	27.4	27.1	2177	4	AAI59411	Aai59411 Human pol
	32	27.4	27.1	7640	4	AAK91239	Aak91239 Human dig
	33	27.4	27.1	8764	4	AAK91240	Aak91240 Human dig
	34	27.2	26.9	326	4	ABA06675	Aba06675 Human cDN
	35	27.2	26.9	326	4	AAS31605	Aas31605 cDNA enco
	36	27.2	26.9	326	6	ABV84012	Abv84012 Human pol
	37	27.2	26.9	570	9	ADD47395	Add47395 Rat gene
	38	27.2	26.9	993	7	ACA35947	Aca35947 Prokaryot
	39	27	26.7	615	6	ABQ53011	Abq53011 Oligonucl
c	40	27	26.7	615	6	ABQ53010	Abq53010 Oligonucl
	41	27	26.7	952	2	AAX04400	Aax04400 Human sec
c	42	27	26.7	2003	2	AAV84571	Aav84571 Human sec
c	43	27	26.7	2003	4	ABA83354	Aba83354 Human sec
c	44	27	26.7	2003	8	ACH04855	Ach04855 Novel hum
c	45	27	26.7	2003	8	ACD44665	Acd44665 Human cDN

ALIGNMENTS

RESULT 1
ABT11173
ID ABT11173 standard; DNA; 168174 BP.
XX
AC ABT11173;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 63.
XX
KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200262825-A2.
XX
PD 15-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-US003546.
XX
PR 08-FEB-2001; 2001US-0267515P.
PR 21-AUG-2001; 2001US-0314248P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes G, Meyer J;
XX
DR WPI; 2002-627522/67.
XX
PT New isolated nucleic acid molecule with an allelic variant of a
PT polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT prognosticating disorders associated with an aberrant inflammatory
PT response such as asthma.
XX
PS Disclosure; Fig 4; 290pp; English.
XX
CC The invention relates to an isolated human nucleic acid molecule
CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC selected from any of 3, 20 or 21 base pair sequences, given in the
CC specification, or their complement. The compositions and methods of the

CC present invention are useful for diagnosing and/or prognosing disorders
CC associated with an aberrant inflammatory response such as asthma,
CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC disease, atopic and contact dermatitis. The nucleic acid molecules can
CC also be useful for identifying an individual amongst other individuals
CC from the same species for use in forensic medicine and paternity testing.
CC This polynucleotide sequence represents DNA relating to the human 5-
CC lipoxygenase (5-LO) gene of the invention
XX
SQ Sequence 168174 BP; 46808 A; 36442 C; 36942 G; 46474 T; 0 U; 1508 Other;

Query Match 88.1%; Score 89; DB 6; Length 168174;
Best Local Similarity 99.0%; Pred. No. 7.6e-17;
Matches 100; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     167319 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGC-CGCCCAGGAGCGCGCGAAACCTTCT 167377

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              |||||||||||||||||||||||||||||||||||||||||
Db     167378 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 167418
```

RESULT 2
ABT11114
ID ABT11114 standard; DNA; 168273 BP.
XX
AC ABT11114;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 2.
XX
KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200262825-A2.
XX
PD 15-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-US003546.
XX
PR 08-FEB-2001; 2001US-0267515P.
PR 21-AUG-2001; 2001US-0314248P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes G, Meyer J;
XX
DR WPI; 2002-627522/67.
XX
PT New isolated nucleic acid molecule with an allelic variant of a
PT polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT prognosticating disorders associated with an aberrant inflammatory
PT response such as asthma.
XX
PS Disclosure; Fig 2; 290pp; English.
XX
CC The invention relates to an isolated human nucleic acid molecule
CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC selected from any of 3, 20 or 21 base pair sequences, given in the
CC specification, or their complement. The compositions and methods of the
CC present invention are useful for diagnosing and/or prognosing disorders
CC associated with an aberrant inflammatory response such as asthma,
CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC disease, atopic and contact dermatitis. The nucleic acid molecules can
CC also be useful for identifying an individual amongst other individuals
CC from the same species for use in forensic medicine and paternity testing.
CC This polynucleotide sequence represents DNA relating to the human 5-
CC lipoxygenase (5-LO) gene of the invention
XX
SQ Sequence 168273 BP; 46834 A; 36467 C; 36966 G; 46498 T; 0 U; 1508 Other;

Query Match 88.1%; Score 89; DB 6; Length 168273;
Best Local Similarity 99.0%; Pred. No. 7.6e-17;
Matches 100; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     167418 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGC-CGCCCAGGAGCGCGCGAAACCTTCT 167476

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              |||||||||||||||||||||||||||||||||||||||||
Db     167477 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 167517
```

RESULT 3
ABL32117/c
ID ABL32117 standard; DNA; 13249 BP.
XX
AC ABL32117;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 90.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 90; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

Query Match 54.5%; Score 55; DB 6; Length 13249;
Best Local Similarity 78.8%; Pred. No. 7.8e-07;
Matches 78; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
Db 9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy 63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
Db 9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992

RESULT 4
ABK31177/c
ID ABK31177 standard; DNA; 13249 BP.
XX
AC ABK31177;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #10.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007472.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction.
XX
PS Claim 1; SEQ ID NO 20; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC the cytosine methylation state (CpG islands) of these genes, and a method
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC genes associated with signal transduction. The genomic DNA can be
CC obtained from cells or cellular components which contain DNA, e.g. cell
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC brain, heart, prostate, lung, breast or liver, histologic object slides,
CC and all their possible combinations. The sequences of the invention are
CC useful for the diagnosis and therapy of diseases associated with signal
CC transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC chemically pretreated genomic DNA sequences of different genes associated
CC with signal transduction, or their complementary sequences. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

Query Match 54.5%; Score 55; DB 6; Length 13249;
Best Local Similarity 78.8%; Pred. No. 7.8e-07;
Matches 78; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
Db 9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy 63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
Db 9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992

RESULT 5
ABL70132/c
ID ABL70132 standard; DNA; 13249 BP.
XX
AC ABL70132;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence complementary to#11.
XX
KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007471.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 22; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

```
  Query Match             54.5%;  Score 55;  DB 6;  Length 13249;
  Best Local Similarity   78.8%;  Pred. No. 7.8e-07;
  Matches   78;  Conservative    0;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||     |||  ||||| |||||  || | ||||||  | ||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              ||||||||||  ||||| |||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992


RESULT 6
ABL32116
ID   ABL32116 standard; DNA; 13249 BP.
XX
AC   ABL32116;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 89.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 89; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity   71.0%;  Pred. No. 0.0026;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||  |||||||  |||     || |||||||||  ||   ||||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  |||  |||||  || |||||||  ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 7
ABK31176
ID   ABK31176 standard; DNA; 13249 BP.
XX
AC   ABK31176;
XX
DT   23-APR-2002  (first entry)
XX
DE   Signal transduction associated gene modified DNA #10.
XX
KW   Human; signal transduction associated gene; cytosine methylation state;
KW   CpG island; signal transduction associated disease; solid tumour; cancer;
KW   antitumour; cytostatic; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007472.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction.
XX
PS   Claim 1; SEQ ID NO 19; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC   Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC   the cytosine methylation state (CpG islands) of these genes, and a method
CC   for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC   genes associated with signal transduction. The genomic DNA can be
CC   obtained from cells or cellular components which contain DNA, e.g. cell
CC   lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC   tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC   brain, heart, prostate, lung, breast or liver, histologic object slides,
CC   and all their possible combinations. The sequences of the invention are
CC   useful for the diagnosis and therapy of diseases associated with signal
CC   transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC   chemically pretreated genomic DNA sequences of different genes associated
CC   with signal transduction, or their complementary sequences. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity   71.0%;  Pred. No. 0.0026;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||  |||||||  |||     || |||||||||  ||   ||||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217
```

```
Qy          61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
               | |   ||  ||| |||||  || ||||||| ||||||
Db        4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 8
ABL70131
ID   ABL70131 standard; DNA; 13249 BP.
XX
AC   ABL70131;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated cell signalling DNA sequence#11.
XX
KW   Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW   tumour; cytostatic; ds.
XX
OS   Unidentified.
XX
PN   WO200202807-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007471.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154758/20.
XX
PT   Nucleic acid, useful for diagnosis and therapy of diseases associated
PT   with cell signaling e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with cell signaling.
XX
PS   Claim 1; SEQ ID NO 21; 24pp + Sequence Listing; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of at least
CC   18 bases of a segment of chemically pretreated DNA of genes associated
CC   with cell signalling. The activity of the modified sequences of the
CC   invention may be described as cytostatic. The object of the invention is
CC   to provide the chemically modified DNA of genes associated with cell
CC   signalling, as well as oligonucleotides and/or PNA-oligomers for
CC   detecting cytosine methylations, as well as a method which is
CC   particularly suitable for the diagnosis and/or therapy of genetic and
CC   epigenetic parameters of genes associated with cell signalling. The
CC   chemically modified DNA provided by the invention is useful for diagnosis
CC   and therapy of diseases such as solid tumours and cancer. The sequences
CC   given in records ABL70111-ABL70626 represent chemically pre-treated
CC   genomic DNA's of genes associated with cell signalling. Note: The
CC   sequence data for this patent is not represented in the printed
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match           42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity   71.0%;  Pred. No. 0.0026;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy           1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
               ||||  ||||||| |||    || |||||||||  ||   ||||||||||||||  || |
Db        4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy          61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
               | |   ||  ||| |||||  || ||||||| ||||||
Db        4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257
```

```
RESULT 9
ABK84481/c
ID   ABK84481 standard; cDNA; 44147 BP.
XX
AC   ABK84481;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #1052.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200228999-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US030821.
XX
PR   03-OCT-2000; 2000US-0237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y,  Weissman SM,  Yamaga S,  Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression of
PT   genes associated with granulocyte activation, which serves as diagnostic
PT   markers that is useful for monitoring disease states and drug toxicity.
XX
PS   Claim 1; SEQ ID NO 1052; 114pp; English.
XX
CC   The invention relates to detecting (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing the
CC   expression level to an expression level in an unactivated GC, where
CC   differential expression of Gs is indicative of GCA. Also included are
CC   modulating (M2) GA by contacting GC with an agent that alters the
CC   expression of at least one gene in Gs; (2) screening (M3) for an agent
CC   capable of modulating GCA or an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease using the gene expression
CC   profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease, by detecting the level of
CC   expression in a sample of the tissue of gene(s) from Gs, where the level
CC   of expression of the gene is indicative of inflammation; (4) treating
CC   (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease, by contacting a tissue having inflammation with an
CC   agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC   is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC   for screening an agent capable of modulating GCA preferably in an
CC   inflammation in a tissue; M4 is useful for detecting an inflammation
CC   (especially chronic) in a tissue, an allergic response in a subject,
CC   exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC   psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC   cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC   respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC   disease, ulcerative colitis, periodontal disease; also bacterial
CC   infection, viral infection, parasitic infection, protozoal infection,
CC   fungal infection and M5 is useful for treating one of the above
CC   conditions. The present sequence represents a gene differentially
```

CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 44147 BP; 9813 A; 11709 C; 11789 G; 10833 T; 0 U; 3 Other;

Query Match 29.7%; Score 30; DB 6; Length 44147;
Best Local Similarity 59.3%; Pred. No. 26;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
Db 29358 GGGGCGGGGCTGCAGCCGCCTGGGCTGTTCCAACGAGCGCGCGCCGCCCCCGCCGCGCCC 29299

Qy 69 TTCCAGGCATTTGCCCGCCGCGATTC 94
Db 29298 CGCGCCGCCCTTGCTAGCCGAGCCTC 29273

RESULT 10
ADD14691/c
ID ADD14691 standard; cDNA; 44147 BP.
XX
AC ADD14691;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polynucleotide SEQ ID NO:85.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR WPI; 2003-636735/60.
DR P-PSDB; ADD14095.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
PS Claim 2; SEQ ID NO 85; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 44147 BP; 9813 A; 11709 C; 11789 G; 10833 T; 0 U; 3 Other;

Query Match 29.7%; Score 30; DB 9; Length 44147;
Best Local Similarity 59.3%; Pred. No. 26;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
Db 29358 GGGGCGGGGCTGCAGCCGCCTGGGCTGTTCCAACGAGCGCGCGCCGCCCCCGCCGCGCCC 29299

Qy 69 TTCCAGGCATTTGCCCGCCGCGATTC 94
Db 29298 CGCGCCGCCCTTGCTAGCCGAGCCTC 29273

RESULT 11
AAL50815/c
ID AAL50815 standard; DNA; 33737 BP.
XX
AC AAL50815;
XX
DT 30-JAN-2003 (first entry)
XX
DE Human cancer status prediction method-related DNA sequence #7.
XX
KW Human; gene therapy; cancer status prediction; cancer; ds;
KW cancer malignancy evaluation; drug design; antisense nucleic acid.
XX
OS Homo sapiens.
XX
PN WO200272828-A1.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-JP002153.
XX
PR 14-MAR-2001; 2001JP-00073063.
PR 06-APR-2001; 2001JP-00108503.
PR 02-AUG-2001; 2001JP-00234807.
XX
PA (DNAC-) DNA CHIP RES INC.
PA (HISF) HITACHI SOFTWARE ENG CO LTD.
XX
PI Kato K, Iwao K, Noguchi S, Matoba R;
XX
DR WPI; 2002-713517/77.
XX
PT Computer-aided statistical method for predicting cancer, applicable in
PT gene therapy for evaluating cancer malignancy with data for use in drug
PT design.
XX
PS Disclosure; Page 132-152; 182pp; Japanese.
XX
CC The invention comprises a method for predicting cancer status. The method
CC involves: measuring expression doses of genes obtained from specimens;
CC selecting at least one gene as the gene for an assay; using the
CC measurement results on expression doses of the selected genes for
CC multivariate analysis; and classifying the specimens in analogous groups
CC with results of the multivariate analysis on expression patterns of the
CC genes. The method of the invention is useful for predicting cancer, which
CC is applicable in gene therapy for evaluating cancer malignancy with data

```
CC   for use in drug design (e.g. antisense nucleic acids for use in gene
CC   therapy to treat cancer). The present DNA sequence represents a human
CC   nucleic acid of the invention
XX
SQ   Sequence 33737 BP; 8084 A; 9254 C; 8733 G; 7663 T; 0 U; 3 Other;

  Query Match             29.1%;  Score 29.4;  DB 6;  Length 33737;
  Best Local Similarity   66.7%;  Pred. No. 37;
  Matches   42;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         10 GGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCT 69
              || ||||| | ||| | | || ||||| |||||  || |  |  | |  |||||| |||
Db       8876 GGCTCAGGGCTCAGTCCCTGGCAGCGCCCCCAGACGCCCAGGGCAGCCCCTCCACCTCCT 8817

Qy         70 TCC 72
              ||
Db       8816 ACC 8814


RESULT 12
ABZ59005
ID   ABZ59005 standard; DNA; 956 BP.
XX
AC   ABZ59005;
XX
DT   28-APR-2003  (first entry)
XX
DE   Human oncosuppressive gene (DRAGO) fragment.
XX
KW   Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;
KW   DRAGO; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003006498-A2.
XX
PD   23-JAN-2003.
XX
PF   09-JUL-2002; 2002WO-EP007625.
XX
PR   10-JUL-2001; 2001IT-MI001465.
XX
PA   (NOVU-) NOVUSPHARMA SPA.
XX
PI   Broggini M,  D'incalci M;
XX
DR   WPI; 2003-221715/21.
XX
PT   New oncosuppressive polypeptide, useful for preparing a medicament for
PT   treating tumors.
XX
PS   Claim 3; Page 42; 42pp; English.
XX
CC   The invention relates to oncosuppressive polypeptides and encoding
CC   polynucleotides. The oncosuppressive gene is involved in apoptotic
CC   process and is regulated by p53 and p73. The oncosuppressive
CC   polynucleotides are useful for preparing a medicament for treating
CC   tumour. The present sequence represents a human oncosuppressive gene
CC   (DRAGO) fragment regulating the expression of the polypeptide
XX
SQ   Sequence 956 BP; 104 A; 371 C; 353 G; 128 T; 0 U; 0 Other;

  Query Match             28.7%;  Score 29;  DB 7;  Length 956;
  Best Local Similarity   58.8%;  Pred. No. 33;
  Matches   50;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || ||   || | | | || ||| || ||  || | |  ||| |||   |  | | |
Db         40 CCCGGCGGGGCGGTCGCGGCGGCGCGACGCCGGCGCCGTTGCGTGCGCGCCTGGCCCTGC 99

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGC 89
                || |||| | |  || |||||| |
Db        100 CCCTTTCCCGCCCCTTCCCCGCCCC 124


RESULT 13
ABZ52368/c
ID   ABZ52368 standard; cDNA; 990 BP.
XX
AC   ABZ52368;
XX
DT   28-MAR-2003  (first entry)
XX
DE   Aspergillus oryzae polynucleotide SEQ ID NO 1481.
XX
KW   Aspergillus oryzae; fermentation; fungus; industrial; EST;
KW   expressed sequence tag; gene; ss.
XX
OS   Aspergillus oryzae.
XX
PN   WO200279476-A1.
XX
PD   10-OCT-2002.
XX
PF   22-MAR-2002; 2002WO-IB000890.
XX
PR   30-MAR-2001; 2001JP-00098371.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (NARE-) NAT RES INST BREWING.
PA   (NORQ ) NAT FOOD RES INST MIN AGRIC.
XX
PI   Machida M,  Akita O,  Kashiwagi Y,  Kitamoto K,  Horiuchi H;
PI   Takeuchi M,  Kobayashi T,  Kitamoto N,  Gomi K,  Abe K;
XX
DR   WPI; 2003-046817/04.
XX
PT   Detection of expression of specific Aspergillus genes for monitoring the
PT   fermentation and growth conditions of the fungus, using DNA probes.
XX
PS   Claim 1; SEQ ID NO 1481; 48pp + Sequence Listing; Japanese.
XX
CC   The invention relates to a polynucleotide having any of 6006 specific
CC   sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC   specific culture conditions including one or more of eutrophic,
CC   oligotrophic, solid, early germination, alkaline, high temperature, low
CC   temperature or maltose culture or polynucleotides stringently hybridising
CC   to these sequences. The polynucleotides are useful for monitoring the
CC   progress of fermentation and the growth conditions of a fungus,
CC   especially of Aspergillus oryzae which is widely used in industrial
CC   fermentation. Also monitoring for fungal contamination. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 990 BP; 247 A; 262 C; 243 G; 235 T; 0 U; 3 Other;

  Query Match             28.7%;  Score 29;  DB 7;  Length 990;
  Best Local Similarity   58.8%;  Pred. No. 33;
  Matches   50;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy         17 GCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTTCCAGGC 76
              | |||||  ||| ||||||  || | | |||  || ||||||  |  |||   || | | |
Db        467 GTCCCAGGCGCCGGAAGCAGGCACCGCAGCCTGCCAAACCTGCCGCACCATCTGCGACGG 408

Qy         77 ATTTGCCCGCCGCGATTCAGAGAGC 101
              | |  | | |   ||     || ||
Db        407 AGTCCCACCCAAAGAAGGCAAGCGC 383


RESULT 14
ABZ58996
ID   ABZ58996 standard; DNA; 3761 BP.
XX
AC   ABZ58996;
XX
```

DT 28-APR-2003 (first entry)
XX
DE Human oncosuppressive gene (DRAGO) fragment.
XX
KW Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;
KW DRAGO; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003006498-A2.
XX
PD 23-JAN-2003.
XX
PF 09-JUL-2002; 2002WO-EP007625.
XX
PR 10-JUL-2001; 2001IT-MI001465.
XX
PA (NOVU-) NOVUSPHARMA SPA.
XX
PI Broggini M, D'incalci M;
XX
DR WPI; 2003-221715/21.
XX
PT New oncosuppressive polypeptide, useful for preparing a medicament for
PT treating tumors.
XX
PS Claim 3; Page 37-38; 42pp; English.
XX
CC The invention relates to oncosuppressive polypeptides and encoding
CC polynucleotides. The oncosuppressive gene is involved in apoptotic
CC process and is regulated by p53 and p73. The oncosuppressive
CC polynucleotides are useful for preparing a medicament for treating
CC tumour. The present sequence represents a human oncosuppressive gene
CC (DRAGO) fragment
XX
SQ Sequence 3761 BP; 800 A; 968 C; 991 G; 1002 T; 0 U; 0 Other;

Query Match 28.7%; Score 29; DB 7; Length 3761;
Best Local Similarity 58.8%; Pred. No. 38;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || ||   || | | || ||| || ||  || | |  ||| |||   |   | | |
Db         45 CCCGGCGGGGCGGTCGCGGCGGCGCGACGCCGGCGCCGTTGCGTGCGCGCCTGGCCCTGC 104

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGC 89
                || |||| | |  || |||||| |
Db        105 CCCTTTCCCGCCCCTTCCCCGCCCC 129
```

RESULT 15
ABZ58995
ID ABZ58995 standard; DNA; 10735 BP.
XX
AC ABZ58995;
XX
DT 28-APR-2003 (first entry)
XX
DE Human oncosuppressive gene (DRAGO) fragment.
XX
KW Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;
KW DRAGO; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003006498-A2.
XX
PD 23-JAN-2003.
XX
PF 09-JUL-2002; 2002WO-EP007625.
XX
PR 10-JUL-2001; 2001IT-MI001465.
XX
PA (NOVU-) NOVUSPHARMA SPA.
XX
PI Broggini M, D'incalci M;
XX
DR WPI; 2003-221715/21.
XX
PT New oncosuppressive polypeptide, useful for preparing a medicament for
PT treating tumors.
XX
PS Claim 3; Page 34-37; 42pp; English.
XX
CC The invention relates to oncosuppressive polypeptides and encoding
CC polynucleotides. The oncosuppressive gene is involved in apoptotic
CC process and is regulated by p53 and p73. The oncosuppressive
CC polynucleotides are useful for preparing a medicament for treating
CC tumour. The present sequence represents a human oncosuppressive gene
CC (DRAGO) fragment, located upstream of the first exon
XX
SQ Sequence 10735 BP; 2720 A; 2485 C; 2630 G; 2900 T; 0 U; 0 Other;

Query Match 28.7%; Score 29; DB 7; Length 10735;
Best Local Similarity 58.8%; Pred. No. 43;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
Qy          5 CCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCAC 64
              || || ||   || | | || ||| || ||  || | |  ||| |||   |   | | |
Db       7019 CCCGGCGGGGCGGTCGCGGCGGCGCGACGCCGGCGCCGTTGCGTGCGCGCCTGGCCCTGC 7078

Qy         65 ACCCTTCCAGGCATTTGCCCGCCGC 89
                || |||| | |  || |||||| |
Db       7079 CCCTTTCCCGCCCCTTCCCCGCCCC 7103
```

Search completed: May 7, 2004, 13:50:23
Job time : 209.335 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:24:59 ; Search time 1052.91 Seconds
(without alignments)
4157.648 Million cell updates/sec

Title: US-10-071-411A-1_COPY_950_1050
Perfect score: 101
Sequence: 1 ggaacctgggggtcaggccc.....cccgccgcgattcagagagc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940541

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		89	88.1	129266	9	AL731567	AL731567 Human DNA
2		89	88.1	160654	2	AC011879	AC011879 Homo sapi
3	c	55	54.5	13249	6	AX344173	AX344173 Sequence
4	c	55	54.5	13249	6	AX345019	AX345019 Sequence
5	c	55	54.5	13249	6	AX348564	AX348564 Sequence
6		43.2	42.8	13249	6	AX344172	AX344172 Sequence
7		43.2	42.8	13249	6	AX345018	AX345018 Sequence
8		43.2	42.8	13249	6	AX348563	AX348563 Sequence
9	c	35	34.7	228272	2	AC098124	AC098124 Rattus no
10		35	34.7	258227	2	AC106376	AC106376 Rattus no
11		35	34.7	307042	2	AC130592	AC130592 Rattus no
12	c	34.4	34.1	199471	2	AC093452	AC093452 Mus muscu
13		32.2	31.9	253540	2	AC093960	AC093960 Rattus no
14	c	32.2	31.9	257344	2	AC111855	AC111855 Rattus no
15		32	31.7	12557	1	AE001980	AE001980 Deinococc
16		31.6	31.3	97517	2	AC091670	AC091670 Oryza sat
17	c	31.6	31.3	133889	8	AC133334	AC133334 Oryza sat
18	c	31.4	31.1	128428	2	AC144889	AC144889 Bos tauru
19		31.4	31.1	251269	2	AC145193	AC145193 Gallus ga
20	c	31	30.7	90175	9	HSDJ858B6	AL118511 Human DNA
21	c	30.8	30.5	325530	2	AC125699	AC125699 Rattus no
22		30.4	30.1	171123	2	AC142009	AC142009 Rattus no
23	c	30.4	30.1	181782	2	AC116970	AC116970 Rattus no
24		30.4	30.1	189029	10	AL831754	AL831754 Mouse DNA
25	c	30.4	30.1	189126	9	AC073094	AC073094 Homo sapi
26	c	30.2	29.9	124413	9	AL354919	AL354919 Human DNA
27		30.2	29.9	183248	2	AC129965	AC129965 Sus scrof
28	c	30.2	29.9	191117	9	AC068629	AC068629 Homo sapi
29	c	30	29.7	44147	9	HSU93305	U93305 Homo sapien
30	c	30	29.7	140335	9	AF235097	AF235097 Homo sapi
31	c	30	29.7	183996	9	AF196779	AF196779 Homo sapi
32	c	30	29.7	225637	2	AC125294	AC125294 Rattus no
33	c	29.8	29.5	884	1	AF239240	AF239240 Ralstonia
34	c	29.8	29.5	2597	9	AB067772	AB067772 Homo sapi
35		29.8	29.5	3025	9	BC048428	BC048428 Homo sapi
36		29.8	29.5	11925	1	AE005111	AE005111 Halobacte
37	c	29.8	29.5	66563	2	AC095759_3	Continuation (4 of
38		29.8	29.5	110000	2	AC095759_1	Continuation (2 of
39		29.8	29.5	140092	9	AL162615	AL162615 Human DNA
40		29.8	29.5	193050	1	AL646062	AL646062 Ralstonia
41	c	29.8	29.5	196050	1	AL646058	AL646058 Ralstonia
42		29.8	29.5	215050	1	AL646084	AL646084 Ralstonia
43		29.8	29.5	216050	1	AL646076	AL646076 Ralstonia
44		29.8	29.5	220050	1	AL646074	AL646074 Ralstonia
45	c	29.8	29.5	231260	2	AL160172	AL160172 Homo sapi

ALIGNMENTS

RESULT 1
AL731567
LOCUS AL731567 129266 bp DNA linear PRI 20-JUN-2002
DEFINITION Human DNA sequence from clone RP11-67C2 on chromosome 10, complete sequence.
ACCESSION AL731567 AC010865
VERSION AL731567.6 GI:21537524
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 129266)
AUTHORS Whitehead,S.
TITLE Direct Submission

```
  JOURNAL   Submitted (31-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 21, 2002 this sequence version replaced gi:21213582.
            Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver
            Street, Waltham, MA 02453, USA
            http://www.genomecorp.com
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-67C2 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .129266
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-67C2"
                     /clone_lib="RPCI-11.1"
ORIGIN

  Query Match             88.1%;  Score 89;  DB 9;  Length 129266;
  Best Local Similarity   99.0%;  Pred. No. 1.7e-13;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      33687 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGC-CGCCCAGGAGCGCGCGAAACCTTCT 33745

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              |||||||||||||||||||||||||||||||||||||||||
Db      33746 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 33786


RESULT 2
AC011879
LOCUS       AC011879               160654 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-16P14, WORKING DRAFT SEQUENCE, 30 unordered
            pieces.
ACCESSION   AC011879
VERSION     AC011879.3  GI:7239554
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-16P14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 14, 2000 this sequence version replaced gi:6524208.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3606
                Center clone name: 16_P_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 111055 bases at least Q40
                Consensus quality: 135066 bases at least Q30
                Consensus quality: 147921 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 157754; sum-of-contigs
                Quality coverage: 2.9 in Q20 bases; agarose-fp
                Quality coverage: 3.0 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 30 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      151: contig of 151 bp in length
            *      152      251: gap of 100 bp
            *      252     1760: contig of 1509 bp in length
            *     1761     1860: gap of 100 bp
            *     1861     3069: contig of 1209 bp in length
            *     3070     3169: gap of 100 bp
            *     3170     4720: contig of 1551 bp in length
            *     4721     4820: gap of 100 bp
            *     4821     6174: contig of 1354 bp in length
            *     6175     6274: gap of 100 bp
            *     6275     7417: contig of 1143 bp in length
            *     7418     7517: gap of 100 bp
            *     7518     9158: contig of 1641 bp in length
            *     9159     9258: gap of 100 bp
            *     9259    10865: contig of 1607 bp in length
            *    10866    10965: gap of 100 bp
            *    10966    12859: contig of 1894 bp in length
            *    12860    12959: gap of 100 bp
            *    12960    15671: contig of 2712 bp in length
            *    15672    15771: gap of 100 bp
            *    15772    18082: contig of 2311 bp in length
            *    18083    18182: gap of 100 bp
            *    18183    20523: contig of 2341 bp in length
            *    20524    20623: gap of 100 bp
```

```
            *    20624    22903: contig of 2280 bp in length
            *    22904    23003: gap of 100 bp
            *    23004    23671: contig of 668 bp in length
            *    23672    23771: gap of 100 bp
            *    23772    25541: contig of 1770 bp in length
            *    25542    25641: gap of 100 bp
            *    25642    28323: contig of 2682 bp in length
            *    28324    28423: gap of 100 bp
            *    28424    31498: contig of 3075 bp in length
            *    31499    31598: gap of 100 bp
            *    31599    36626: contig of 5028 bp in length
            *    36627    36726: gap of 100 bp
            *    36727    42109: contig of 5383 bp in length
            *    42110    42209: gap of 100 bp
            *    42210    48339: contig of 6130 bp in length
            *    48340    48439: gap of 100 bp
            *    48440    55333: contig of 6894 bp in length
            *    55334    55433: gap of 100 bp
            *    55434    63594: contig of 8161 bp in length
            *    63595    63694: gap of 100 bp
            *    63695    73943: contig of 10249 bp in length
            *    73944    74043: gap of 100 bp
            *    74044    83665: contig of 9622 bp in length
            *    83666    83765: gap of 100 bp
            *    83766    95322: contig of 11557 bp in length
            *    95323    95422: gap of 100 bp
            *    95423   108403: contig of 12981 bp in length
            *   108404   108503: gap of 100 bp
            *   108504   120220: contig of 11717 bp in length
            *   120221   120320: gap of 100 bp
            *   120321   132958: contig of 12638 bp in length
            *   132959   133058: gap of 100 bp
            *   133059   145697: contig of 12639 bp in length
            *   145698   145797: gap of 100 bp
            *   145798   160654: contig of 14857 bp in length.
FEATURES             Location/Qualifiers
     source          1. .160654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-16P14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .151
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    252. .1760
                     /note="assembly_fragment"
     misc_feature    1861. .3069
                     /note="assembly_fragment"
     misc_feature    3170. .4720
                     /note="assembly_fragment"
     misc_feature    4821. .6174
                     /note="assembly_fragment"
     misc_feature    6275. .7417
                     /note="assembly_fragment"
     misc_feature    7518. .9158
                     /note="assembly_fragment"
     misc_feature    9259. .10865
                     /note="assembly_fragment"
     misc_feature    10966. .12859
                     /note="assembly_fragment"
     misc_feature    12960. .15671
                     /note="assembly_fragment"
     misc_feature    15772. .18082
                     /note="assembly_fragment"
     misc_feature    18183. .20523
                     /note="assembly_fragment"
     misc_feature    20624. .22903
                     /note="assembly_fragment"
     misc_feature    23004. .23671
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    23772. .25541
                     /note="assembly_fragment"
     misc_feature    25642. .28323
                     /note="assembly_fragment"
     misc_feature    28424. .31498
                     /note="assembly_fragment"
     misc_feature    31599. .36626
                     /note="assembly_fragment"
     misc_feature    36727. .42109
                     /note="assembly_fragment"
     misc_feature    42210. .48339
                     /note="assembly_fragment"
     misc_feature    48440. .55333
                     /note="assembly_fragment"
     misc_feature    55434. .63594
                     /note="assembly_fragment"
     misc_feature    63695. .73943
                     /note="assembly_fragment"
     misc_feature    74044. .83665
                     /note="assembly_fragment"
     misc_feature    83766. .95322
                     /note="assembly_fragment"
     misc_feature    95423. .108403
                     /note="assembly_fragment"
     misc_feature    108504. .120220
                     /note="assembly_fragment"
     misc_feature    120321. .132958
                     /note="assembly_fragment"
     misc_feature    133059. .145697
                     /note="assembly_fragment"
     misc_feature    145798. .160654
                     /note="assembly_fragment"
ORIGIN

  Query Match             88.1%;  Score 89;  DB 2;  Length 160654;
  Best Local Similarity   99.0%;  Pred. No. 1.6e-13;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      40764 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGC-CGCCCAGGAGCGCGCGAAACCTTCT 40822

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              |||||||||||||||||||||||||||||||||||||||||
Db      40823 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 40863


RESULT 3
AX344173/c
LOCUS       AX344173               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 20 from Patent WO0200926.
ACCESSION   AX344173
VERSION     AX344173.1  GI:18492061
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 20 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             54.5%;  Score 55;  DB 6;  Length 13249;
  Best Local Similarity   78.8%;  Pred. No. 0.00017;
```

```
  Matches   78;  Conservative    0;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||     |||  |||||  |||||  || |  ||||||  | ||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              ||||||||||||  |||||  ||||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992


RESULT 4
AX345019/c
LOCUS       AX345019               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 90 from Patent WO0200928.
ACCESSION   AX345019
VERSION     AX345019.1  GI:18492905
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 90 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            54.5%;  Score 55;  DB 6;  Length 13249;
  Best Local Similarity  78.8%;  Pred. No. 0.00017;
  Matches   78;  Conservative    0;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||     |||  |||||  |||||  || |  ||||||  | ||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              ||||||||||||  |||||  ||||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992


RESULT 5
AX348564/c
LOCUS       AX348564               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 22 from Patent WO0202807.
ACCESSION   AX348564
VERSION     AX348564.1  GI:18614599
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 22 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            54.5%;  Score 55;  DB 6;  Length 13249;
  Best Local Similarity  78.8%;  Pred. No. 0.00017;
  Matches   78;  Conservative    0;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||     |||  |||||  |||||  || |  ||||||  | ||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              ||||||||||||  |||||  ||||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992


RESULT 6
AX344172
LOCUS       AX344172               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 19 from Patent WO0200926.
ACCESSION   AX344172
VERSION     AX344172.1  GI:18492060
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 19 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity  71.0%;  Pred. No. 0.2;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||  |||||||  |||     || |||||||||  ||   ||||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  ||| |||||  || |||||||  ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 7
AX345018
LOCUS       AX345018               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 89 from Patent WO0200928.
ACCESSION   AX345018
VERSION     AX345018.1  GI:18492904
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 89 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity  71.0%;  Pred. No. 0.2;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          3 AACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCC 62
              |||||     |||  |||||  |||||  || |  ||||||  | ||||||||||||||||
Db       9089 AACCTAAAAATCAAACCCCAACCGCGAAAAAC-CGCCCAAAAACGCGCGAAACCTTCTCC 9031

Qy         63 ACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
              ||||||||||||  |||||  ||||||||||||||| | | |
Db       9030 ACACCCTTCCAAACATTTACCCGCCGCGATTCAAAAAAC 8992


RESULT 6
AX344172
LOCUS       AX344172               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 19 from Patent WO0200926.
ACCESSION   AX344172
VERSION     AX344172.1  GI:18492060
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 19 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity  71.0%;  Pred. No. 0.2;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||||  |||||||  |||     || |||||||||  ||   ||||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  ||| |||||  || |||||||  ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 7
AX345018
LOCUS       AX345018               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 89 from Patent WO0200928.
ACCESSION   AX345018
VERSION     AX345018.1  GI:18492904
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 89 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match            42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity  71.0%;  Pred. No. 0.2;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;
```

```
Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              |||| ||||||| |||    || |||||||||  ||   |||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  ||| |||||  || |||||| ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 8
AX348563
LOCUS       AX348563               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 21 from Patent WO0202807.
ACCESSION   AX348563
VERSION     AX348563.1  GI:18614598
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 21 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             42.8%;  Score 43.2;  DB 6;  Length 13249;
  Best Local Similarity   71.0%;  Pred. No. 0.2;
  Matches   71;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              |||| ||||||| |||    || |||||||||  ||   |||||||||||||  || |
Db       4159 GGAATTTGGGGGTTAGGTTTTAGTCGCGGGAAG-TCGTTTAGGAGCGCGCGAAATTTTTT 4217

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAGAGAG 100
                | |   ||  ||| |||||  || |||||| ||||||
Db       4218 TTATATTTTTTTAGGTATTTGTTCGTCGCGATTTAGAGAG 4257


RESULT 9
AC098124/c
LOCUS       AC098124               228272 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-134B20, WORKING DRAFT SEQUENCE.
ACCESSION   AC098124
VERSION     AC098124.6  GI:30522711
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 228272)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 228272)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 228272)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:22855450.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GGLR
                Center clone name: CH230-134B20
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 212229 bases at least Q40
                Consensus quality: 213662 bases at least Q30
```

```
            Consensus quality: 214533 bases at least Q20
            Estimated insert size: 221650; sum-of-contigs estimation
            Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
          ------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 1 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *       1   228272: contig of 228272 bp in length.
FEATURES             Location/Qualifiers
     source          1. .228272
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-134B20"
     misc_feature    1. .1792
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    5581. .7407
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    7458. .8579
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    complement(9189. .9393)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH348064"
ORIGIN

  Query Match             34.7%;  Score 35;  DB 2;  Length 228272;
  Best Local Similarity   61.5%;  Pred. No. 22;
  Matches   56;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              | ||||  |      || ||| ||| ||||| |||| ||||||  | |||    |   |
Db      26554 GAAACCGAGCACCGCGGGCCCCGCCACGGGACGCGCTCCCAGGCCCCCGCCCTGCGCACC 26495

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGA 91
              |||| ||||||| ||     || ||| || |
Db      26494 CCACCCCCTTCCCGGGCCGCGCACGCAGCCA 26464


RESULT 10
AC106376
LOCUS       AC106376              258227 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-115H19, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC106376
VERSION     AC106376.5  GI:30581262
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 258227)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 258227)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 258227)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:25007772.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
```

```
            ---------- Project Information
                Center project name: GKTX
                Center clone name: CH230-115H19
            ---------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 217062 bases at least Q40
                Consensus quality: 220975 bases at least Q30
                Consensus quality: 223890 bases at least Q20
                Estimated insert size: 227972; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   255801: contig of 255801 bp in length
            *  255802   255901: gap of unknown length
            *  255902   257111: contig of 1210 bp in length
            *  257112   257211: gap of unknown length
            *  257212   258227: contig of 1016 bp in length.
FEATURES             Location/Qualifiers
     source          1. .258227
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-115H19"
     misc_feature    1. .2031
                     /note="wgs_contig"
     misc_feature    194330. .195710
                     /note="wgs_contig"
ORIGIN

  Query Match             34.7%;  Score 35;  DB 2;  Length 258227;
  Best Local Similarity   61.5%;  Pred. No. 21;
  Matches   56;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              | ||||  |      || ||| ||| ||||| |||| ||||||  | |||    |   |
Db      20616 GAAACCGAGCACCGCGGGCCCCGCCACGGGACGCGCTCCCAGGCCCCCGCCCTGCGCACC 20675

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGA 91
              |||| ||||||| ||     || ||| || |
Db      20676 CCACCCCCTTCCCGGGCCGCGCACGCAGCCA 20706


RESULT 11
AC130592
LOCUS       AC130592               307042 bp    DNA     linear   HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-506G20, WORKING DRAFT SEQUENCE.
ACCESSION   AC130592
VERSION     AC130592.3  GI:24635281
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 307042)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 307042)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 307042)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 6, 2002 this sequence version replaced gi:23664960.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
```

```
            Center project name: KCCD
            Center clone name: CH230-506G20
          ---------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 162473 bases at least Q40
            Consensus quality: 164097 bases at least Q30
            Consensus quality: 165166 bases at least Q20
            Estimated insert size: 165502; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
          ----------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 1 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *       1   307042: contig of 307042 bp in length.
FEATURES             Location/Qualifiers
     source          1. .307042
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-506G20"
     misc_feature    1. .2685
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    complement(6035. .6919)
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:RXBUF46TJB"
     misc_feature    complement(6035. .6875)
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:RXBUF46TJ"
     misc_feature    167089. .167976
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:RXBUF46TVB"
     misc_feature    169057. .170819
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    173435. .175226
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    305571. .307042
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

  Query Match             34.7%;  Score 35;  DB 2;  Length 307042;
  Best Local Similarity   61.5%;  Pred. No. 21;
  Matches   56;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              | ||||  |      || ||| ||| ||||| |||| ||||||  | |||    |   |
Db     149502 GAAACCGAGCACCGCGGGCCCCGCCACGGGACGCGCTCCCAGGCCCCCGCCCTGCGCACC 149561

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGA 91
              |||| ||||||| ||     || ||| || |
Db     149562 CCACCCCCTTCCCGGGCCGCGCACGCAGCCA 149592


RESULT 12
AC093452/c
LOCUS       AC093452               199471 bp    DNA     linear   HTG 19-DEC-2003
DEFINITION  Mus musculus clone rp23-458j13 map 1 strain C57BL/6J, WORKING DRAFT
            SEQUENCE, 8 ordered pieces.
ACCESSION   AC093452
VERSION     AC093452.15  GI:38638756
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 199471)
  AUTHORS   Li,J., Yang,L. and Roe,B.A.
  TITLE     Mus musculus BAC Clone rp23-458j13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199471)
  AUTHORS   Li,J., Yang,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 199471)
  AUTHORS   Li,J., Yang,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Dec 3, 2003 this sequence version replaced gi:38604127.
            ---------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1    39305: contig of 39305 bp in length
            *   39306    39405: gap of unknown length
            *   39406    55773: contig of 16368 bp in length
            *   55774    55873: gap of unknown length
            *   55874    72673: contig of 16800 bp in length
            *   72674    72773: gap of unknown length
            *   72774   158934: contig of 86161 bp in length
            *  158935   159034: gap of unknown length
            *  159035   179496: contig of 20462 bp in length
            *  179497   179596: gap of unknown length
            *  179597   188163: contig of 8567 bp in length
            *  188164   188263: gap of unknown length
            *  188264   196368: contig of 8105 bp in length
            *  196369   196468: gap of unknown length
            *  196469   199471: contig of 3003 bp in length.
FEATURES             Location/Qualifiers
     source          1. .199471
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="1"
                     /clone="rp23-458j13"
                     /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN

  Query Match             34.1%;  Score 34.4;  DB 2;  Length 199471;
  Best Local Similarity   65.8%;  Pred. No. 31;
  Matches   50;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy         16 GGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCCTTCCAGG 75
              || ||| ||| ||||| |||| |||||| | |||    |   | |||| |||||||| ||
Db     52658 GGGCCCCGCCACGGGACGCGCTCCCAGGCCCCCGCCCTGCGCACCCCACCCCCTTCCCGG 52599
```

```
Qy          76 CATTTGCCCGCCGCGA 91
                  || ||| || |
Db       52598 GCCGCGCACGCAGCCA 52583


RESULT 13
AC093960
LOCUS       AC093960               253540 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-4G13, WORKING DRAFT SEQUENCE, 3
            unordered pieces.
ACCESSION   AC093960
VERSION     AC093960.9  GI:30466917
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 253540)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 253540)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 253540)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:24818990.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GAQO
                Center clone name: CH230-4G13
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 237398 bases at least Q40
                Consensus quality: 238410 bases at least Q30
                Consensus quality: 238873 bases at least Q20
                Estimated insert size: 246935; sum-of-contigs estimation
                Quality coverage: 10x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   245384: contig of 245384 bp in length
            *    245385   245484: gap of unknown length
            *    245485   246668: contig of 1184 bp in length
            *    246669   246768: gap of unknown length
            *    246769   253540: contig of 6772 bp in length.
FEATURES             Location/Qualifiers
     source          1. .253540
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-4G13"
     misc_feature    1. .1009
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    1060. .3204
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    complement(4133. .4983)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH306359"
ORIGIN
```

```
  Query Match             31.9%;  Score 32.2;  DB 2;  Length 253540;
  Best Local Similarity   59.1%;  Pred. No. 1.2e+02;
  Matches   55;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              |||| || | | |||  | |||| |   | ||  | | |||| || || |   ||   ||
Db     120760 GGGGACATGGCTCAGTTGGGGGAGGGCTGGCCTAGCGTGCGCCAAGCCATGGTCATCTCC 120819

Qy         69 TTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
                |  | |||   ||||| | | | ||| ||||
Db     120820 AGCACGACATAAACCCGCTGAGGTACAGCGAGC 120852


RESULT 14
AC111855/c
LOCUS       AC111855               257344 bp    DNA     linear   HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-196H16, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC111855
VERSION     AC111855.4  GI:25006939
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 257344)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 257344)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 257344)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 15, 2002 this sequence version replaced gi:23270243.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GOJH
                Center clone name: CH230-196H16
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 238203 bases at least Q40
                Consensus quality: 241708 bases at least Q30
                Consensus quality: 244072 bases at least Q20
                Estimated insert size: 245085; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   228515: contig of 228515 bp in length
            *    228516   228615: gap of unknown length
            *    228616   256091: contig of 27476 bp in length
            *    256092   256191: gap of unknown length
            *    256192   257344: contig of 1153 bp in length.
FEATURES             Location/Qualifiers
     source          1. .257344
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-196H16"
     misc_feature    1. .1227
                     /note="wgs_contig"
     misc_feature    228616. .229887
                     /note="wgs_contig"
     misc_feature    232589. .236151
                     /note="wgs_contig"
```

```
ORIGIN

  Query Match            31.9%;  Score 32.2;  DB 2;  Length 257344;
  Best Local Similarity  59.1%;  Pred. No. 1.2e+02;
  Matches   55;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          9 GGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCTCCACACCC 68
              |||| || | | ||| | |||| |   | ||  | | |||| || || |   ||   ||
Db     174799 GGGGACATGGCTCAGTTGGGGGAGGGCTGGCCTAGCGTGCGCCAAGCCATGGTCATCTCC 174740

Qy         69 TTCCAGGCATTTGCCCGCCGCGATTCAGAGAGC 101
                |  | |||   ||||| | | | ||| ||||
Db     174739 AGCACGACATAAACCCGCTGAGGTACAGCGAGC 174707


RESULT 15
AE001980
LOCUS       AE001980               12557 bp    DNA     linear   BCT 22-NOV-1999
DEFINITION  Deinococcus radiodurans R1 section 117 of 229 of the complete
            chromosome 1.
ACCESSION   AE001980 AE000513
VERSION     AE001980.1  GI:6459086
KEYWORDS    .
SOURCE      Deinococcus radiodurans R1
  ORGANISM  Deinococcus radiodurans R1
            Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
            Deinococcaceae; Deinococcus.
REFERENCE   1  (bases 1 to 12557)
  AUTHORS   White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D.,
            Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L.,
            Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M.,
            Vamathevan,J.J., Lam,P., McDonald,L., Utterback,T., Zalewski,C.,
            Makarova,K.S., Aravind,L., Daly,M.J., Fraser,C.M. et al.
  TITLE     Genome sequence of the radioresistant bacterium Deinococcus
            radiodurans R1
  JOURNAL   Science 286 (5444), 1571-1577 (1999)
  MEDLINE   20036896
   PUBMED   10567266
REFERENCE   2  (bases 1 to 12557)
  AUTHORS   White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D.,
            Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L.,
            Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M.,
            Vamathevan,J.J., Lam,P., McDonald,L., Utterback,T., Zalewski,C.,
            Makarova,K.S., Aravind,L., Daly,M.J., Minton,K.W.,
            Fleischmann,R.D., Ketchum,K.A., Nelson,K.E., Salzberg,S.,
            Smith,H.O., Venter,J.C. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1999) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .12557
                     /organism="Deinococcus radiodurans R1"
                     /mol_type="genomic DNA"
                     /strain="R1"
                     /db_xref="taxon:243230"
                     /chromosome="1"
     gene            complement(84. .1334)
                     /gene="DR1334"
     CDS             complement(84. .1334)
                     /gene="DR1334"
                     /note="similar to PID:2257467 percent identity: 100.00;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="transposase, putative"
                     /protein_id="AAF10906.1"
                     /db_xref="GI:6459087"
                     /translation="MNTPLPMGLMAILQYVLSAVPLRKTQRNFLTVLLSVFLAVPGRL
                     NVLNLSRYAACSESTIRRWLHRSDPGAIPWGAVHRATVSTAIESGLISPLCVLAIDAS
                     FHRKSGQHTAHLGSFWNGCAARTERGIEQSCCALIDVQHRQAFTVDVRQTRTGSEAPS
                     RLEQAADQLDDVLLDLQTVPRLDLAAVVADGNYAKESMVETVTGHGLPFISRFPRNAN
                     LKYLYTGEHPRRRGRPKKFDGKVDFSDLQRFDLVSETSTERVWTQVVWSVQWAREVRA
                     VVIQQVGKKGQVTGYAVLFSTAVTMPAHEVIALYRSRFEIELIFRDAKQFLGGQDVQL
                     RSQPGIEAHWNVVLLTLNLCRLEALRAAEGGQNLVFSLEDMKRRAYNALLAQVILSKL
                     DLSARFEELEHLPSSPLNFGLKAA"
     gene            complement(1432. .4668)
                     /gene="DR1335"
     CDS             complement(1432. .4668)
                     /gene="DR1335"
                     /note="similar to GB:X75439 GB:X59478 SP:P41368 PID:581558
                     percent identity: 59.12; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="isoleucyl-tRNA synthetase"
                     /protein_id="AAF10907.1"
                     /db_xref="GI:6459088"
                     /translation="MTKRTFAPVPTQPNFRQLEADILQKWKDEQVFEQTQTRPAPAGE
                     FVFYEGPPTANGKPALHHVLARSFKDLFPRFKVMQGHHVTRKGGWDTHGLPVEISVEK
                     KLGWLGRNHGASREELEEFNRLCRTSVWETIQDWNELTERMGYWLDLGDPYITYQNSY
                     VESVWNLLRRLHEKGLVAQDYKVVPLSPRISSTLSRAELGEVDSYRMVDDPSVYVRLP
                     IVWDTLPERAHAALSSLSGEQRQGLSLLVWTTTPWTLPSNTLAAVNPDLDYVVADSPS
                     GRVIVAEGAVERLSALHKDAAPLEVLARFKGRDLEWVEYEPPFPEVASQLGVVSELHE
                     RRDGKPVLHFVVMADFVSDVDGSGVAHEAPAYGAEDLEVARAYGVPLMFGVDDHGILQ
                     VTHEQGKFFKDADKGLIADMKARGLMFWAGTLKHRYPFHDRTGDPILYFAKKGWYIRT
                     SQVAGEMLAQNEKINWVPGNIKHGRFGNWLEGNVDWAISRERYWGTPLPFWQSESGQL
                     RVIGSVAELSELAGRDLSDLDLHRPYIDDITFTLDGEEYRRVPEVLDVWFDSGSMPYA
                     QWGLLLNEQGEAVRGAEQFAKHYPADYICEAIDQTRGWFYSLHAISTMLYDQPAYKNV
                     ICLGHIVDEKGLKMSKSKGNVVAPLPLFDQYGADSVRWYMFMASDPGDQKRFSERLVA
                     EAQRSYVNTLWNVYSFFVLYANLDQPDLAAAPAVAERPEMDRWLLARLEETVRDVTAA
                     LESYDARSGGRALEGFVDDLSNWYVRRSRSRFWGEGGTVDTAAYATLHEALLVVSQLT
                     APFTPFLADALYRNLSGEESSVHLTPWPTVRAERLDRKLTADMAAVMKVVELGRAVRG
                     AHNLKTRQPLAGVQVRAASPEALDALKRSQTQIMEELNVKAVTFLEGDTDLVQYSLRP
                     NLPVVGKQYGKQLPVLKKALTEADARAVATAVQAGQGFSVQADGVTFDLTPGSVLVDA
                     KAPEGVAAAEDAGYLVAFDTALTPELVREGLARDLVRAIQEARKAAGFEVQDRIALAL
                     ELDGEALEAAQAWQDFIAGEVLAEQVAYGSGEGFRAEVEGGAVTLKKL"
     gene            4693. .6354
                     /gene="DR1336"
     CDS             4693. .6354
                     /gene="DR1336"
                     /note="similar to GB:M76990 SP:P07775 PID:141751 percent
                     identity: 59.18; identified by sequence similarity;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="transport protein, putative"
                     /protein_id="AAF10908.1"
                     /db_xref="GI:6459089"
                     /translation="MMVDGAFVRAEEAVSLNRLPSPISKNPAPAAHSDLRLSTCRDAL
                     ELAAWYHPTSPPIFGGLIPRPSAGGRPGRSSGPQAVLLPAREGIFAARSGAVRLSPSL
                     TRSCPRVRGYWPHVRLGCPGMLRRGGARGQCAIQRRTARKRGAAFRQPPGPLPAPAQP
                     CTVGGGRVPAPYFPRAARRGLSDWGAGGFAAAHSSPAQGASALASMYLLLGVMGMTLC
                     WRYRVPVVLGWNTPGLAVLIASAGRFTAAEVVGALLVSGAVLALVGLSRAFDWLARTL
                     PLPLALALLAGVLLPFVLRGLAAIPQAPALTLPTLLAYLLARVFAPRWAVPLALVTGV
                     TVAALTGQLHLSALGHTSLLARPEWITPTFSSQALLTLALPCTILALTSQHLAGMAVL
                     HTHGYRQVPPALLVTVTGAASALAAPLGSVTMNMGAITAAICMGPDVHPDPNRHYVAG
                     LACAGGYLVLGLCAGALLGLAGSFPAPLMQLLAGLALLGSVLSGLQNALSDPRWREGA
                     LLAFACTAAGFSFLNIGGALWGLLLGWGSAALAQWAQREKGGGLA"
     gene            6341. .7201
                     /gene="DR1337"
     CDS             6341. .7201
                     /gene="DR1337"
                     /note="similar to GB:AE000657 percent identity: 74.31;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="transaldolase"
                     /protein_id="AAF10909.1"
                     /db_xref="GI:6459090"
                     /translation="MGWPEGRAPVAPEQGAPGPFFSDRGGVCWRHPAQVRQGKSGAGP
                     AVPLRRTHMEFFIDTAIVDEIREINAWGVLSGVTTNPSLIVASGRDFKEVIGEIAELV
                     PGGAISAEVTALDAEGMIKEGKDVAAWHKQVVVKLPLTPAGLQACKALTSEGIKTNVT
                     LCFSVPQALLAARAGATYVSPFVGRVDDTGWDGSELIRQIKEAFVLGDIQTKVLAASI
                     RHPQHVVQAALAGADVATIPYKVFTQMVKHPLTQAGLDSFMADWAKRQGASPETPPSE
                     AGTNPVKKEG"
     gene            7198. .8478
```

```
                     /gene="DR1338"
     CDS             7198. .8478
                     /gene="DR1338"
                     /note="similar to GB:L27276 PID:454814 SP:P52153 percent
                     identity: 99.05; identified by sequence similarity;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="transcription termination factor Rho"
                     /protein_id="AAF10910.1"
                     /db_xref="GI:6459091"
                     /translation="MTVTEVAPQALPFQELQEKILPELHLLAAGLGIENYRKLKKDAL
                     ALAIMEKQADAEGQSLARGYLDITSDGYGFLQADLLDPASRSVLVTAGVIKQYHLRTG
                     DEVIGRARKPRENERYGSLVRVEAVNGLDPEAARQRPRFDDLTPTFPDQQLVLEDPST
                     DDGLSLRVVDLLVPIGRGQRALIVAPPKAGKTTLLKKIANSITKNYPDVTVMVLLVDE
                     RPEEVTDFRESVQGAQVIASTFDEPPQHHVRVAEFVHERARRIVEEGGHVVILLDSIT
                     RLARANNLVTPPTGRTLSGGLDSNALHWPKRFLGAARNIREGGSLTILATALVETGSR
                     MDDVIFEEFKGTGNAELVLSRRLEERRIFPALDILKSGTRREELLLQPEVLKKMWLLR
                     KVISDMDPADAMEMLLGRMGKTRNNVEFLAALAG"
     gene            complement(8620. .9354)
                     /gene="DR1339"
     CDS             complement(8620. .9354)
                     /gene="DR1339"
                     /note="similar to SP:P19583 GB:X59403 PID:40500
                     PID:1124842 percent identity: 68.80; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="triosephosphate isomerase"
                     /protein_id="AAF10911.1"
                     /db_xref="GI:6459092"
                     /translation="MQTLLALNWKMNKTPTEARSWAEELTTKYAPAEGVDLAVLAPAL
                     DLSALAANLPAGIAFGGQDVSAHESGAYTGEISAAMLKDAGASCVVVGHSERREYHDE
                     SDAXVAAKARQAQANGLLPIVCVGENLDVRERGEHVPQTLAQLRGSLEGVGADVVVAY
                     EPVWAIGTGKTATADDAEELAAAIRGALREQYGARAEGIRVLYGGSVKPENIAEICGK
                     PNVNGALVGGASLKVPDVLGMLDALR"
     gene            complement(9416. .9778)
                     /gene="DR1340"
     CDS             complement(9416. .9778)
                     /gene="DR1340"
                     /note="similar to SP:P71604 PID:1552589 GB:AL123456
                     percent identity: 44.54; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF10912.1"
                     /db_xref="GI:6459093"
                     /translation="MNINDARDLITRLFTIIDSADYDRLGDVFAADAVYERPGYEPLQ
                     GLPRIEQFYRHERVIGSGRHTVEDVTCSETGSAVSFGVFRGTSRAGEALEERFADVYR
                     VQDGKIVQRTTYFFRPAI"
     gene            complement(9775. .10209)
                     /gene="DR1341"
     CDS             complement(9775. .10209)
                     /gene="DR1341"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF10915.1"
                     /db_xref="GI:6459096"
                     /translation="MSEPDLVPELAVADLATSLGFWVNLLGFGLKYERPEEGFAYLTL
                     GNAHSMLDQIGQSRTWETAPLQRPLGRGINFEISVSDLDVPLARLMAADWPLFLAPEE
                     KWYRAGDHETGVRQFLVQDPDGYLVRLSISLGQCPLLQEATR"
     gene            complement(10206. .11441)
                     /gene="DR1342"
     CDS             complement(10206. .11441)
                     /gene="DR1342"
                     /note="similar to SP:P09403 GB:X12464 PID:48249 percent
                     identity: 76.94; identified by sequence similarity;
                     putative"

Query Match          31.7%;  Score 32;  DB 1;  Length 12557;
```

```
  Best Local Similarity   58.3%;  Pred. No. 1.7e+02;
  Matches   56;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 GGAACCTGGGGGTCAGGCCCCAGCCGCGGGAAGCGCGCCCAGGAGCGCGCGAAACCTTCT 60
              ||  || ||||  | || |||  |  ||| |  ||| | | |||||| |  | | | |
Db       5074 GGGGCCAGGGGTCAATGCGCCATTCAGCGGAGGACCGCACGGAAGCGCGGGGCAGCGTTT 5133

Qy         61 CCACACCCTTCCAGGCATTTGCCCGCCGCGATTCAG 96
                ||| ||  || | |  |||||||  ||   |||
Db       5134 AGACAGCCCCCCGGCCCGCTGCCCGCGCCGGCCCAG 5169


Search completed: May  7, 2004, 14:31:04
Job time : 1059.25 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:34:53 ; Search time 1484.07 Seconds
(without alignments)
1629.864 Million cell updates/sec

Title: US-10-071-411A-1_COPY_500_580
Perfect score: 81
Sequence: 1 tcaagaaagtgaaaacacaa.....ttagagacttctatccagga 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	50.6	62.5	832	28	BZ850187	BZ850187 CH240_195
c	2	48.6	60.0	603	14	CF362037	CF362037 828421 MA
	3	48.6	60.0	652	29	CE678956	CE678956 tigr-gss-
c	4	48.6	60.0	657	14	CF364190	CF364190 834119 MA
c	5	48.6	60.0	701	14	CF181752	CF181752 818731 MA
	6	47.4	58.5	360	9	AA569625	AA569625 nm38g05.s
c	7	47.4	58.5	361	10	BE142248	BE142248 CM1-HT014
c	8	47.4	58.5	422	29	CE268863	CE268863 tigr-gss-
c	9	47.4	58.5	643	28	AQ373217	AQ373217 RPCI11-14
c	10	47.4	58.5	657	13	BX480547	BX480547 DKFZp686C
	11	47.2	58.3	585	12	BI057063	BI057063 PM2-GN050
	12	47.2	58.3	625	12	BI057060	BI057060 PM2-GN050
	13	47	58.0	752	9	AU122280	AU122280 AU122280
	14	46.4	57.3	462	28	AQ270907	AQ270907 HS_2059_A
	15	46.4	57.3	569	28	AQ385233	AQ385233 RPCI11-14
c	16	46.4	57.3	948	29	ECA542812	AJ542812 Equus cab
c	17	46	56.8	387	9	AA366464	AA366464 EST77411
c	18	46	56.8	475	29	CE053095	CE053095 tigr-gss-
c	19	46	56.8	527	28	AQ664300	AQ664300 HS_5493_B
c	20	46	56.8	704	29	AG185330	AG185330 Pan trogl
c	21	45.8	56.5	209	14	F35156	F35156 HSPD30917 H
	22	45.8	56.5	228	9	AA805911	AA805911 oc13c02.s
	23	45.8	56.5	232	10	AW129543	AW129543 xe18f10.x
c	24	45.8	56.5	368	9	AA631500	AA631500 np83h06.s
c	25	45.8	56.5	383	14	F29519	F29519 HSPD19437 H
	26	45.8	56.5	404	9	AA551394	AA551394 nk71h06.s
	27	45.8	56.5	422	12	BQ004338	BQ004338 UI-H-EI0-
	28	45.8	56.5	445	9	AI049591	AI049591 an34d09.x
c	29	45.8	56.5	453	9	AA506540	AA506540 ng13h10.s
c	30	45.8	56.5	483	10	AW853687	AW853687 RC3-CT025
	31	45.8	56.5	483	29	CE535406	CE535406 tigr-gss-
	32	45.8	56.5	493	9	AA602341	AA602341 no89h01.s
	33	45.8	56.5	497	9	AA602613	AA602613 no91c07.s
	34	45.8	56.5	519	9	AA604512	AA604512 no73c09.s
	35	45.8	56.5	536	9	AA604920	AA604920 no81b04.s
c	36	45.8	56.5	538	29	CE155843	CE155843 tigr-gss-
c	37	45.8	56.5	546	28	AQ591742	AQ591742 HS_5428_A
c	38	45.8	56.5	565	13	BX641858	BX641858 DKFZp686N
	39	45.8	56.5	573	9	AU145980	AU145980 AU145980
c	40	45.8	56.5	588	29	CE624617	CE624617 tigr-gss-
	41	45.8	56.5	591	9	AI963674	AI963674 wr64e08.x
	42	45.8	56.5	615	29	CE414829	CE414829 tigr-gss-
c	43	45.8	56.5	629	13	BX492832	BX492832 DKFZp781L
	44	45.8	56.5	631	29	CE603572	CE603572 tigr-gss-
	45	45.8	56.5	678	13	BU688601	BU688601 UI-CF-DU1

ALIGNMENTS

RESULT 1
BZ850187
LOCUS BZ850187 832 bp DNA linear GSS 18-MAR-2003
DEFINITION CH240_195F24.TVB CHORI-240 Bos taurus genomic clone CH240_195F24, genomic survey sequence.
ACCESSION BZ850187
VERSION BZ850187.1 GI:29077588
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 832)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P., Crawford,A.M. and McEwan,J.C.
TITLE Bovine BAC End Sequences from Library CHORI-240
JOURNAL Unpublished (2003)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by AgResearch Ltd., New Zealand and The Institute of Genomic Research (TIGR), USA.
Plate: 195 row: F column: 24
Seq primer: T7
Class: BAC ends.

```
FEATURES             Location/Qualifiers
     source          1. .832
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_195F24"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             62.5%;  Score 50.6;  DB 28;  Length 832;
  Best Local Similarity   76.5%;  Pred. No. 0.00061;
  Matches   62;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||||||| |||| ||||||| || ||  |  |||||| | ||  | ||||||||| |
Db         16 TCAAGAAAGTAAAAAGACAACCCACAAAATGAGGAAAAATATTTGCCAATCATGTATCTG 75

Qy         61 ATTAGAGACTTCTATCCAGGA 81
                || ||||||| |||  || |
Db         76 ATAAGAGACTTTTATAGAGAA 96


RESULT 2
CF362037/c
LOCUS       CF362037                603 bp    mRNA    linear   EST 25-AUG-2003
DEFINITION  828421 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF362037
VERSION     CF362037.1  GI:34161295
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
            Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     A second set of porcine ESTs from a pooled-tissue normalized
            library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: SRG8018  row: J  column: 16
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .603
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 3PIG"
                     /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                     Library made with RNA pooled from multiple tissues
                     including brain, liver, muscle, placenta/endometrium,
                     ovary, testes, and bone marrow."
ORIGIN

  Query Match             60.0%;  Score 48.6;  DB 14;  Length 603;
  Best Local Similarity   75.9%;  Pred. No. 0.0022;
  Matches   60;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              |||||||||| || ||||||| |||||  |   ||||| | ||  | |||  |||| |||
Db        165 AAGAAAGTGAGAAGACAACCCACAGAATGAGAGAAAATATTTGCCAATCACATATCTGAT 106

Qy         63 TAGAGACTTCTATCCAGGA 81
                |||||||| ||||||| |
Db        105 AAGAGACTTGTATCCAGAA 87


RESULT 3
CE678956
LOCUS       CE678956                652 bp    DNA     linear   GSS 29-SEP-2003
DEFINITION  tigr-gss-dog-17000313980922 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE678956
VERSION     CE678956.1  GI:36997956
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             60.0%;  Score 48.6;  DB 29;  Length 652;
  Best Local Similarity   75.9%;  Pred. No. 0.0022;
  Matches   60;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||||||| |||  ||||||| ||||| ||   ||||| | || || |||| |||
Db         33 TCAAGAAAGTTAAAGAACAACCCACAGAATCAGAGAAAATATTTGCAAATCATATATTTA 92

Qy         61 ATTAGAGACTTCTATCCAG 79
                || || |||||  |||||||
Db         93 ATAAGGGACTTGTATCCAG 111


RESULT 4
```

```
CF364190/c
LOCUS       CF364190                657 bp    mRNA    linear   EST 25-AUG-2003
DEFINITION  834119 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF364190
VERSION     CF364190.1  GI:34165561
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
            Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     A second set of porcine ESTs from a pooled-tissue normalized
            library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: SRG8020  row: P  column: 19
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 3PIG"
                     /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                     Library made with RNA pooled from multiple tissues
                     including brain, liver, muscle, placenta/endometrium,
                     ovary, testes, and bone marrow."
ORIGIN

  Query Match             60.0%;  Score 48.6;  DB 14;  Length 657;
  Best Local Similarity   75.9%;  Pred. No. 0.0022;
  Matches   60;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              |||||||||| || ||||||| |||||  |   ||||| | ||  | |||  |||| |||
Db        165 AAGAAAGTGAGAAGACAACCCACAGAATGAGAGAAAATATTTGCCAATCACATATCTGAT 106

Qy         63 TAGAGACTTCTATCCAGGA 81
              |||||||| ||||||| |
Db        105 AAGAGACTTGTATCCAGAA 87


RESULT 5
CF181752/c
LOCUS       CF181752                701 bp    mRNA    linear   EST 28-JUL-2003
DEFINITION  818731 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF181752
VERSION     CF181752.1  GI:33293528
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
            Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     A second set of porcine ESTs from a pooled-tissue normalized
            library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: SRG8013  row: J  column: 7
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 3PIG"
                     /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                     Library made with RNA pooled from multiple tissues
                     including brain, liver, muscle, placenta/endometrium,
                     ovary, testes, and bone marrow."
ORIGIN

  Query Match             60.0%;  Score 48.6;  DB 14;  Length 701;
  Best Local Similarity   75.9%;  Pred. No. 0.0022;
  Matches   60;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              |||||||||| || ||||||| |||||  |   ||||| | ||  | |||  |||| |||
Db        165 AAGAAAGTGAGAAGACAACCCACAGAATGAGAGAAAATATTTGCCAATCACATATCTGAT 106

Qy         63 TAGAGACTTCTATCCAGGA 81
              |||||||| ||||||| |
Db        105 AAGAGACTTGTATCCAGAA 87


RESULT 6
AA569625
LOCUS       AA569625                360 bp    mRNA    linear   EST 09-SEP-1997
DEFINITION  nm38g05.s1 NCI_CGAP_Pr4.1 Homo sapiens cDNA clone IMAGE:1062488
            similar to contains L1.b1 L1 repetitive element ;, mRNA sequence.
ACCESSION   AA569625
VERSION     AA569625.1  GI:2343605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 360)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 339   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .360
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1062488"
                     /sex="male"
                     /tissue_type="prostatic intraepithelial neoplasia - high
```

grade"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr4.1"
/note="Organ: prostate; Vector: pAMP10; mRNA made from prostatic intraepithelial neoplasia (high-grade), cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. cDNA Library Preparation: David B. Krizman, Ph.D. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center"

ORIGIN

```
  Query Match             58.5%;  Score 47.4;  DB 9;  Length 360;
  Best Local Similarity   74.1%;  Pred. No. 0.005;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||||||||||||| ||||||| || ||   |  |||||  | || || |||| |||| |
Db         99 TCAAGAAAGTGAAAAGACAACCCACAAAATGAGAAAAAACATTTGCAAATCATATATCTG 158

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || |  | ||  ||||||| |
Db        159 ATAAAGGTCTAATATCCAGAA 179
```

RESULT 7
BE142248/c

```
LOCUS       BE142248                361 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  CM1-HT0141-170999-013-h05 HT0141 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE142248
VERSION     BE142248.1  GI:8604969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 361)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM1-HT0141-170
            999-013-h05&t3=1999-09-17&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 361.
FEATURES             Location/Qualifiers
     source          1. .361
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0141"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 - Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
```

ORIGIN

```
  Query Match             58.5%;  Score 47.4;  DB 10;  Length 361;
  Best Local Similarity   74.1%;  Pred. No. 0.005;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              | |||||| |||||| |||||||| || || |   |||||  | |  |  | |||||||  |
Db         81 TGAAGAAAATGAAAAGACAACCCGTAGCAGAACAGAAAATATTTTCACATAATGTATCTG 22

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || |||||||| |||| || |
Db         21 ATGAGAGACTTGTATCTAGTA 1
```

RESULT 8
CE268863/c

```
LOCUS       CE268863                422 bp    DNA     linear   GSS 26-SEP-2003
DEFINITION  tigr-gss-dog-17000359927985 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE268863
VERSION     CE268863.1  GI:35997099
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
```

ORIGIN

```
  Query Match             58.5%;  Score 47.4;  DB 29;  Length 422;
  Best Local Similarity   74.1%;  Pred. No. 0.0049;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| |||||| |||||      || || |||| || |||||||||
Db        229 TCAAAAAAGTGAAAAGACAACCCACAGAATGGGAGAAGATTTCTGCAAATCATGTATCTC 170

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              | || ||||| | || ||||
Db        169 GTAAGGGACTTATTTCTAGGA 149
```

```
RESULT 9
AQ373217/c
LOCUS       AQ373217                643 bp    DNA     linear   GSS 20-MAY-1999
DEFINITION  RPCI11-146M20.TV RPCI-11 Homo sapiens genomic clone RPCI-11-146M20,
            genomic survey sequence.
ACCESSION   AQ373217
VERSION     AQ373217.1  GI:4344240
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-146M20.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7555987"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-146M20"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             58.5%;  Score 47.4;  DB 28;  Length 643;
  Best Local Similarity   74.1%;  Pred. No. 0.0047;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| ||||||||||  |||||| |||||     |   ||| | ||||| |||| |||  |
Db        427 TCAACAAAGTGAAAAAGCAACCCACAGAATGGGTGGGAATATTTGTAAATCATATATTTG 368

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || ||||| ||||||| |
Db        367 ATAAGGGACTTTTATCCAGAA 347


RESULT 10
BX480547/c
LOCUS       BX480547                657 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686C23222_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686C23222 5', mRNA sequence.
ACCESSION   BX480547
VERSION     BX480547.1  GI:31917361
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686C23222) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686C23222"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 657;
  Best Local Similarity   74.1%;  Pred. No. 0.0046;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| ||||||||||  |||||| |||||     |   ||| | ||||| |||| |||  |
Db        266 TCAAAAAAGTGAAAAGACAACTCACAGAAAGGGAGAAAATATTTGCAAATCATGTATCTG 207

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || ||||| |  |||| |
Db        206 ATAAGGGACTTGTGCCCAGAA 186


RESULT 11
BI057063
LOCUS       BI057063                585 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  PM2-GN0500-220201-002-h01 GN0500 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI057063
VERSION     BI057063.1  GI:14464593
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
```

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM2&t2=PM2-GN0500-220201-002-h01&t3=2001-02-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 523.

FEATURES Location/Qualifiers
source 1. .585
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0500"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 58.3%; Score 47.2; DB 12; Length 585;
Best Local Similarity 76.3%; Pred. No. 0.0053;
Matches 58; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy          2 CAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGA 61
              ||||||||   | | ||||||| |||||      |||||  | |||||  ||||||||||||
Db         33 CAAGAAAGCAGAGAGACAACCCACAGAATGGCAGAAAATATTTGTAAATCATGTATCCGA 92

Qy         62 TTAGAGACTTCTATCC 77
              | || ||||| |||||
Db         93 TAAGGGACTTGTATCC 108
```

RESULT 12
BI057060
LOCUS BI057060 625 bp mRNA linear EST 15-JUN-2001
DEFINITION PM2-GN0500-220201-002-b06 GN0500 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI057060
VERSION BI057060.1 GI:14464590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM2&t2=PM2-GN0500-220201-002-b06&t3=2001-02-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 283.

FEATURES Location/Qualifiers
source 1. .625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0500"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 58.3%; Score 47.2; DB 12; Length 625;
Best Local Similarity 76.3%; Pred. No. 0.0053;
Matches 58; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy          2 CAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGA 61
              ||||||||   | | ||||||| |||||      |||||  | |||||  ||||||||||||
Db         33 CAAGAAAGCAGAGAGACAACCCACAGAATGGCAGAAAATATTTGTAAATCATGTATCCGA 92

Qy         62 TTAGAGACTTCTATCC 77
              | || ||||| |||||
Db         93 TAAGGGACTTGTATCC 108
```

RESULT 13
AU122280
LOCUS AU122280 752 bp mRNA linear EST 01-AUG-2002
DEFINITION AU122280 MAMMA1 Homo sapiens cDNA clone MAMMA1002056 5', mRNA sequence.
ACCESSION AU122280
VERSION AU122280.1 GI:10937515
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES Location/Qualifiers
source 1. .752
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1002056"
/tissue_type="mammary gland"
/clone_lib="MAMMA1"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 58.0%; Score 47; DB 9; Length 752;
Best Local Similarity 74.7%; Pred. No. 0.0058;
Matches 59; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              || |||||||||| |||||||  ||||  |  |||||| | |  ||| ||||||||| |||
Db        620 AAAAAAGTGAAAATACAACCCATAGAAAGATAAAAAATATTTTCAAGCCATGTATCTGAT 679

Qy         63 TAGAGACTTCTATCCAGGA 81
                || | ||  ||||||| |
Db        680 AAGGGTCTAGTATCCAGAA 698
```

RESULT 14
AQ270907
LOCUS AQ270907 462 bp DNA linear GSS 03-NOV-1998
DEFINITION HS_2059_A2_A07_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2059 Col=14 Row=A, genomic survey sequence.
ACCESSION AQ270907
VERSION AQ270907.1 GI:3823447
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 462)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2059 row: A column: 14
Class: BAC ends
High quality sequence stop: 462.
FEATURES Location/Qualifiers
source 1. .462
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2059 Col=14 Row=A"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
ORIGIN

Query Match 57.3%; Score 46.4; DB 28; Length 462;
Best Local Similarity 72.8%; Pred. No. 0.0091;
Matches 59; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||| || | |||||| ||| ||| |||||      |||||| | ||||| |||| ||||||
Db         21 TCAGGANAATGAAAAGACACCCCACAGAATGGGAGAAAATATTTGTAAATCATATATCCG 80

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||  ||||||| |
Db         81 ATAAGGGTCTAGTATCCAGAA 101
```

RESULT 15
AQ385233
LOCUS AQ385233 569 bp DNA linear GSS 21-MAY-1999
DEFINITION RPCI11-141M1.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-141M1, genomic survey sequence.
ACCESSION AQ385233
VERSION AQ385233.1 GI:4356256
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPCI11-141M1.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .569
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7554048"
/db_xref="taxon:9606"
/clone="RPCI-11-141M1"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
ORIGIN

Query Match 57.3%; Score 46.4; DB 28; Length 569;
Best Local Similarity 73.8%; Pred. No. 0.0088;
Matches 59; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy          2 CAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGA 61
              |||| || |||| | ||||||| |||||      ||||||   || || |||||||||  ||
Db         37 CAAGGAACTGAAGAGACAACCCACAGAATGGGAAAAAATACTTGCAAATCATGTATCTGA 96

Qy         62 TTAGAGACTTCTATCCAGGA 81
              | |||| ||  ||||||| |
Db         97 TAAGAGTCTAGTATCCAGAA 116
```

Search completed: May 7, 2004, 15:42:30
Job time : 1487.74 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:31:30 ; Search time 167.215 Seconds
(without alignments)
2194.362 Million cell updates/sec

Title: US-10-071-411A-1_COPY_500_580
Perfect score: 81
Sequence: 1 tcaagaaagtgaaaacacaa.....ttagagacttctatccagga 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883171

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	60.4	74.6	13249	15	US-10-311-455-90	Sequence 90, Appl
	2	58.6	72.3	13249	15	US-10-311-455-89	Sequence 89, Appl
c	3	51	63.0	331	13	US-10-085-783A-12549	Sequence 12549, A
c	4	51	63.0	331	16	US-10-242-535A-12549	Sequence 12549, A
c	5	50.2	62.0	72332	12	US-10-052-482-58	Sequence 58, Appl
	6	49	60.5	95683	13	US-10-087-192-160	Sequence 160, App
c	7	47.4	58.5	592	13	US-10-027-632-48192	Sequence 48192, A
c	8	47.4	58.5	592	13	US-10-027-632-48193	Sequence 48193, A
c	9	47.4	58.5	592	13	US-10-027-632-48194	Sequence 48194, A
c	10	47.4	58.5	592	13	US-10-027-632-48195	Sequence 48195, A
c	11	47.4	58.5	592	16	US-10-027-632-48192	Sequence 48192, A
c	12	47.4	58.5	592	16	US-10-027-632-48193	Sequence 48193, A
c	13	47.4	58.5	592	16	US-10-027-632-48194	Sequence 48194, A
c	14	47.4	58.5	592	16	US-10-027-632-48195	Sequence 48195, A
	15	47.4	58.5	638	13	US-10-027-632-78918	Sequence 78918, A
	16	47.4	58.5	638	13	US-10-027-632-78919	Sequence 78919, A
	17	47.4	58.5	638	13	US-10-027-632-78920	Sequence 78920, A
	18	47.4	58.5	638	13	US-10-027-632-78921	Sequence 78921, A
	19	47.4	58.5	638	13	US-10-027-632-314671	Sequence 314671,
	20	47.4	58.5	638	13	US-10-027-632-314672	Sequence 314672,
	21	47.4	58.5	638	13	US-10-027-632-314673	Sequence 314673,
	22	47.4	58.5	638	13	US-10-027-632-314674	Sequence 314674,
	23	47.4	58.5	638	16	US-10-027-632-78918	Sequence 78918, A
	24	47.4	58.5	638	16	US-10-027-632-78919	Sequence 78919, A
	25	47.4	58.5	638	16	US-10-027-632-78920	Sequence 78920, A
	26	47.4	58.5	638	16	US-10-027-632-78921	Sequence 78921, A
	27	47.4	58.5	638	16	US-10-027-632-314671	Sequence 314671,
	28	47.4	58.5	638	16	US-10-027-632-314672	Sequence 314672,
	29	47.4	58.5	638	16	US-10-027-632-314673	Sequence 314673,
	30	47.4	58.5	638	16	US-10-027-632-314674	Sequence 314674,
c	31	47.4	58.5	128034	13	US-10-282-174-186	Sequence 186, App
c	32	47.4	58.5	128034	13	US-10-282-174-187	Sequence 187, App
	33	47.4	58.5	202100	13	US-10-282-174-484	Sequence 484, App
	34	47.4	58.5	2140405	13	US-10-027-632-76212	Sequence 76212, A
	35	47.4	58.5	2140405	16	US-10-027-632-76212	Sequence 76212, A
c	36	45.8	56.5	432	9	US-09-920-455-146	Sequence 146, App
c	37	45.8	56.5	454	13	US-10-085-783A-41212	Sequence 41212, A
c	38	45.8	56.5	454	16	US-10-242-535A-41212	Sequence 41212, A
c	39	45.8	56.5	478	9	US-09-920-455-202	Sequence 202, App
c	40	45.8	56.5	498	10	US-09-814-353-18323	Sequence 18323, A
c	41	45.8	56.5	523	9	US-09-920-455-199	Sequence 199, App
c	42	45.8	56.5	524	9	US-09-920-455-196	Sequence 196, App
c	43	45.8	56.5	535	10	US-09-814-353-18518	Sequence 18518, A
c	44	45.8	56.5	603	10	US-09-814-353-6173	Sequence 6173, Ap
c	45	45.8	56.5	603	10	US-09-814-353-12451	Sequence 12451, A

ALIGNMENTS

RESULT 1
US-10-311-455-90/c
; Sequence 90, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 90
; LENGTH: 13249
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-90

Query Match 74.6%; Score 60.4; DB 15; Length 13249;
Best Local Similarity 85.9%; Pred. No. 6.6e-10;
Matches 67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| ||| | |||||||||||||||| || |||||||||| ||| ||| |||| ||||||
Db       9541 TCAAAAAAATAAAAACACAACCCGCAAAAACAATAAAAATATCTATAAATCATATATCCG 9482

Qy         61 ATTAGAGACTTCTATCCA 78
```

```
                |||| | ||||||||||||
Db       9481 ATTAAAAACTTCTATCCA 9464


RESULT 2
US-10-311-455-89
; Sequence 89, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 89
;   LENGTH: 13249
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-89

  Query Match             72.3%;  Score 58.6;  DB 15;  Length 13249;
  Best Local Similarity   82.7%;  Pred. No. 2.7e-09;
  Matches   67;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              | |||||||||||||| | ||  || ||||| |||||||||| ||||||| |||||| ||
Db       3709 TTAAGAAAGTGAAAATATAATTCGTAGAAGTAATAAAAATGTTTGTAAGTTATGTATTCG 3768

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              |||||||| || |||  ||||
Db       3769 ATTAGAGATTTTTATTTAGGA 3789


RESULT 3
US-10-085-783A-12549/c
; Sequence 12549, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12549
;   LENGTH: 331
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (137)..(137)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-12549

  Query Match             63.0%;  Score 51;  DB 13;  Length 331;
  Best Local Similarity   78.9%;  Pred. No. 2.8e-07;
  Matches   60;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          2 CAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGA 61
              ||||||| ||||||  |||||| |||||      ||||| | ||||| ||||||||||||
Db        152 CAAGAAACTGAAAAGNCAACCCACAGAATGGCAGAAAATATTTGTAAATCATGTATCCGA 93

Qy         62 TTAGAGACTTCTATCC 77
              | || ||||| |||||
Db         92 TAAGGGACTTGTATCC 77


RESULT 4
US-10-242-535A-12549/c
; Sequence 12549, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12549
;   LENGTH: 331
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (137)..(137)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-12549

  Query Match             63.0%;  Score 51;  DB 16;  Length 331;
  Best Local Similarity   78.9%;  Pred. No. 2.8e-07;
  Matches   60;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          2 CAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGA 61
              ||||||| ||||||  |||||| |||||      ||||| | ||||| ||||||||||||
Db        152 CAAGAAACTGAAAAGNCAACCCACAGAATGGCAGAAAATATTTGTAAATCATGTATCCGA 93

Qy         62 TTAGAGACTTCTATCC 77
              | || ||||| |||||
Db         92 TAAGGGACTTGTATCC 77


RESULT 5
US-10-052-482-58/c
; Sequence 58, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
;  APPLICANT: Engelhard, Eric
;  APPLICANT:  Morris, David
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71087/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/10/052,482
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
```

```
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
;   LENGTH: 72332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (70768)..(71491)
;   OTHER INFORMATION: "n" at positions 70768 to 71491 can be any base
US-10-052-482-58

  Query Match             62.0%;  Score 50.2;  DB 12;  Length 72332;
  Best Local Similarity   77.2%;  Pred. No. 3.3e-06;
  Matches   61;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||| ||||||||| ||||||| |||||      ||||| | || ||||||||||  |
Db      36015 TCAAGCAAGTGAAAAGACAACCCACAGAATGGGGGAAAATATTTGCAAGTCATGTATTGG 35956

Qy         61 ATTAGAGACTTCTATCCAG 79
              || || ||||| ||| |||
Db      35955 ATAAGGGACTTGTATGCAG 35937


RESULT 6
US-10-087-192-160
; Sequence 160, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 160
;   LENGTH: 95683
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-160

  Query Match             60.5%;  Score 49;  DB 13;  Length 95683;
  Best Local Similarity   75.3%;  Pred. No. 9.3e-06;
  Matches   61;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||||||||| ||| ||||||| |||||      ||||| | ||| | | |||| || |
Db      14012 TCAAGAAAGTGGAAAGACAACCCACAGAATGGGAGAAAATATTTGTGAATTATGTCTCTG 14071

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || ||||| ||||||| |
Db      14072 ATAAGGGACTTGTATCCAGAA 14092


RESULT 7
US-10-027-632-48192/c
; Sequence 48192, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48192
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48192

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| |||||| |||||  |  |||||| | |  ||  ||||||||  |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||  |||||||  |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 8
US-10-027-632-48193/c
; Sequence 48193, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48193
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48193

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
```

```
                ||||| |||||||||| |||||| |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 9
US-10-027-632-48194/c
; Sequence 48194, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48194
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48194

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| |||||| |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 10
US-10-027-632-48195/c
; Sequence 48195, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48195
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48195

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| |||||| |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 11
US-10-027-632-48192/c
; Sequence 48192, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48192
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48192

  Query Match             58.5%;  Score 47.4;  DB 16;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| |||||| |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261
```

```
RESULT 12
US-10-027-632-48193/c
; Sequence 48193, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48193
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48193

  Query Match             58.5%;  Score 47.4;  DB 16;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||||  |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 13
US-10-027-632-48194/c
; Sequence 48194, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48194
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48194

  Query Match             58.5%;  Score 47.4;  DB 16;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||||  |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 14
US-10-027-632-48195/c
; Sequence 48195, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48195
;   LENGTH: 592
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-48195

  Query Match             58.5%;  Score 47.4;  DB 16;  Length 592;
  Best Local Similarity   74.1%;  Pred. No. 5.6e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||||  |||||  |  ||||||  |  |  ||  |||||||| |
Db        341 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 282

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  |  ||  ||||||| |
Db        281 ATAAGGGTCTAGTATCCAGAA 261


RESULT 15
US-10-027-632-78918
; Sequence 78918, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
```

```
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78918
;   LENGTH: 638
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-78918

  Query Match             58.5%;  Score 47.4;  DB 13;  Length 638;
  Best Local Similarity   74.1%;  Pred. No. 5.8e-06;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||||  |||||  |  |||||| | |  ||  |||||||| |
Db        263 TCAAAAAAGTGAAAATACAACCAACAGAAAGAGAAAAAATATTTTCAAACCATGTATCTG 322

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||  |||||| |
Db        323 ATAAGGGTCTAGTATCCAGAA 343


Search completed: May  7, 2004, 17:35:53
Job time : 170.215 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:35:03 ; Search time 40.6738 Seconds
(without alignments)
1105.159 Million cell updates/sec

Title: US-10-071-411A-1_COPY_500_580
Perfect score: 81
Sequence: 1 tcaagaaagtgaaaacacaa.....ttagagacttctatccagga 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	45.8	56.5	649	3	US-09-040-984-59	Sequence 59, Appl
c	2	45.8	56.5	649	4	US-09-123-912-59	Sequence 59, Appl
c	3	45.8	56.5	649	4	US-09-643-597-59	Sequence 59, Appl
c	4	45.8	56.5	649	4	US-09-480-884A-59	Sequence 59, Appl
c	5	45.8	56.5	649	4	US-09-542-615A-59	Sequence 59, Appl
c	6	45.8	56.5	649	4	US-09-606-421B-59	Sequence 59, Appl
c	7	45.8	56.5	649	4	US-09-221-107-59	Sequence 59, Appl
	8	42.6	52.6	4080	4	US-09-016-434-1342	Sequence 1342, Ap
	9	36.2	44.7	168575	4	US-09-426-290-1	Sequence 1, Appli
c	10	35.8	44.2	63588	4	US-09-873-404-3	Sequence 3, Appli
c	11	35.6	44.0	788	4	US-09-288-143-27	Sequence 27, Appl
	12	34.8	43.0	3001	4	US-09-539-333D-211	Sequence 211, App
	13	34.6	42.7	99916	4	US-09-816-095-3	Sequence 3, Appli
c	14	34.6	42.7	786431	4	US-09-751-389-3	Sequence 3, Appli
c	15	33.8	41.7	523	4	US-09-621-976-1442	Sequence 1442, Ap
c	16	33	40.7	262	4	US-09-621-976-189	Sequence 189, App
	17	33	40.7	359	4	US-09-621-976-9585	Sequence 9585, Ap
c	18	33	40.7	38564	4	US-09-734-673-3	Sequence 3, Appli
	19	33	40.7	392000	4	US-10-027-983-11	Sequence 11, Appl
c	20	32	39.5	10607	1	US-08-078-090-3	Sequence 3, Appli
c	21	31.8	39.3	505	4	US-09-621-976-3878	Sequence 3878, Ap
	22	31.8	39.3	3757	4	US-09-620-312D-106	Sequence 106, App
c	23	31.8	39.3	392000	4	US-10-027-983-11	Sequence 11, Appl
c	24	31.6	39.0	31208	4	US-09-852-067-3	Sequence 3, Appli
c	25	31.6	39.0	98844	4	US-09-791-211-10	Sequence 10, Appl
	26	31.6	39.0	116592	4	US-09-818-512-3	Sequence 3, Appli
	27	31.4	38.8	343	3	US-08-991-789A-72	Sequence 72, Appl
	28	31.4	38.8	343	4	US-09-062-451-72	Sequence 72, Appl
	29	31.4	38.8	343	4	US-09-598-326-72	Sequence 72, Appl
	30	31.4	38.8	343	4	US-09-289-198-72	Sequence 72, Appl
	31	31.4	38.8	343	4	US-09-429-755-72	Sequence 72, Appl
	32	31.4	38.8	15418	4	US-09-783-203-1	Sequence 1, Appli
	33	31.4	38.8	51552	4	US-09-733-294A-30	Sequence 30, Appl
c	34	31.4	38.8	116592	4	US-09-818-512-3	Sequence 3, Appli
c	35	31.4	38.8	193303	4	US-09-497-855A-37	Sequence 37, Appl
c	36	31.4	38.8	193303	4	US-09-497-855A-44	Sequence 44, Appl
c	37	31.2	38.5	6801	4	US-10-204-708-61	Sequence 61, Appl
c	38	31	38.3	402	4	US-09-621-976-1855	Sequence 1855, Ap
c	39	31	38.3	459	4	US-09-621-976-1849	Sequence 1849, Ap
c	40	30.2	37.3	202001	4	US-09-734-674-3	Sequence 3, Appli
c	41	30	37.0	1010	4	US-09-461-325-102	Sequence 102, App
c	42	30	37.0	1010	4	US-10-012-542-102	Sequence 102, App
	43	30	37.0	65042	4	US-09-784-316-3	Sequence 3, Appli
	44	29.8	36.8	369	3	US-08-991-789A-191	Sequence 191, App
	45	29.8	36.8	369	4	US-09-062-451-191	Sequence 191, App

ALIGNMENTS

RESULT 1
US-09-040-984-59/c
; Sequence 59, Application US/09040984
; Patent No. 6210883
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, TongTong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF LUNG CANCER
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,984
; FILING DATE: 18-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-282-6031
; TELEX:
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-040-984-59

Query Match 56.5%; Score 45.8; DB 3; Length 649;
Best Local Similarity 72.8%; Pred. No. 1.6e-06;
Matches 59; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              || | ||||||||||||||||||| |||||      |||||  | ||  ||  ||||||||  |
Db        309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250
```

```
Qy          61 ATTAGAGACTTCTATCCAGGA 81
               || || | ||   |||||| |
Db         249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 2
US-09-123-912-59/c
; Sequence 59, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE:  210121.455C1
;  CURRENT APPLICATION NUMBER: US/09/123,912A
;  CURRENT FILING DATE:  1998-07-27
;  PRIOR APPLICATION NUMBER:  09/040,802
;  PRIOR FILING DATE:  1998-03-18
;  NUMBER OF SEQ ID NOS:  114
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (22)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (190)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (217)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (430)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (433)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (484)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (544)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (550)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (577)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (583)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (594)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-59

  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy           1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
               || | ||||||||||||||||||| |||||      ||||| | || ||  |||||||| |
Db         309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy          61 ATTAGAGACTTCTATCCAGGA 81
               || || | ||   |||||| |
Db         249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 3
US-09-643-597-59/c
; Sequence 59, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(649)
;   OTHER INFORMATION: n = A,T,C or G
US-09-643-597-59

  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy           1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
               || | ||||||||||||||||||| |||||      ||||| | || ||  |||||||| |
Db         309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy          61 ATTAGAGACTTCTATCCAGGA 81
               || || | ||   |||||| |
Db         249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 4
US-09-480-884A-59/c
; Sequence 59, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C6
;  CURRENT APPLICATION NUMBER: US/09/480,884A
;  CURRENT FILING DATE:  2001-08-27
;  NUMBER OF SEQ ID NOS: 330
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(649)
;   OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-59
```

```
  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              || | |||||||||||||||||| |||||      |||||  | || ||  |||||||| |
Db        309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||   |||||| |
Db        249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 5
US-09-542-615A-59/c
; Sequence 59, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C8
;  CURRENT APPLICATION NUMBER: US/09/542,615A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  350
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(649)
;   OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-59

  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              || | |||||||||||||||||| |||||      |||||  | || ||  |||||||| |
Db        309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||   |||||| |
Db        249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 6
US-09-606-421B-59/c
; Sequence 59, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(649)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-59

  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              || | |||||||||||||||||| |||||      |||||  | || ||  |||||||| |
Db        309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||   |||||| |
Db        249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 7
US-09-221-107-59/c
; Sequence 59, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE: 210121.455C2
;  CURRENT APPLICATION NUMBER: US/09/221,107
;  CURRENT FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 161
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
;   LENGTH: 649
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (22)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (190)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (217)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (430)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (433)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (484)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (544)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (550)
;   OTHER INFORMATION: Where n is a, c, g or t
```

```
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (577)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (583)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (594)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-59

  Query Match             56.5%;  Score 45.8;  DB 4;  Length 649;
  Best Local Similarity   72.8%;  Pred. No. 1.6e-06;
  Matches   59;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              || | ||||||||||||||||||| |||||      |||||  | ||  ||  |||||||| |
Db        309 TCGATAAAGTGAAAACACAACCCACAGAATGGGAGAAAATATTTGCAAACCATGTATCTG 250

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  | ||   |||||| |
Db        249 ATAAGGGTCTAGAATCCAGAA 229


RESULT 8
US-09-016-434-1342
; Sequence 1342, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1342:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4080 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g34764
US-09-016-434-1342

  Query Match             52.6%;  Score 42.6;  DB 4;  Length 4080;
  Best Local Similarity   70.4%;  Pred. No. 2.9e-05;
  Matches   57;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||||| ||||||| | |||| ||||||      ||||| | ||  |||||| ||| |
Db       3163 TCAAGAAGGTGAAAAGATAACCTGCAGAATGGGAGAAAATATTTGCGAGTCATATATATG 3222

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || | ||| ||||  | |
Db       3223 ATAAGGGGCTTGTATCTGGAA 3243


RESULT 9
US-09-426-290-1
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
;  APPLICANT: Berglind Ran Olafsdottir
;  APPLICANT: Jeffrey Gulcher
;  TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
;  FILE REFERENCE: 2345.2001-000
;  CURRENT APPLICATION NUMBER: US/09/426,290
;  CURRENT FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 168575
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (21181)...(21403)
;   NAME/KEY: CDS
;   LOCATION: (95252)...(95430)
;   NAME/KEY: CDS
;   LOCATION: (101753)...(101996)
;   NAME/KEY: CDS
;   LOCATION: (110324)...(110439)
;   NAME/KEY: CDS
;   LOCATION: (124058)...(124278)
;   NAME/KEY: CDS
;   LOCATION: (127009)...(127130)
;   NAME/KEY: CDS
;   LOCATION: (128910)...(129139)
US-09-426-290-1

  Query Match             44.7%;  Score 36.2;  DB 4;  Length 168575;
  Best Local Similarity   65.4%;  Pred. No. 0.011;
  Matches   53;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||||   |||||| | ||||  || ||      || |||| || || ||||||||  |
Db      12256 TCAAGAGTATGAAAAGAAAACCTACAAAATAGGAGAACATGTTTGGAAATCATGTATTTG 12315

Qy         61 ATTAGAGACTTCTATCCAGGA 81
                | || || |  ||| ||| |
Db      12316 TTAAGGGATTAGTATACAGAA 12336


RESULT 10
US-09-873-404-3/c
; Sequence 3, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001212-CIP
;  CURRENT APPLICATION NUMBER: US/09/873,404
```

```
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 63588
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(63588)
;   OTHER INFORMATION: n = A,T,C or G
US-09-873-404-3

  Query Match             44.2%;  Score 35.8;  DB 4;  Length 63588;
  Best Local Similarity   65.8%;  Pred. No. 0.011;
  Matches   52;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| | ||| || | ||||||  |||||   |  ||||| | || ||   ||| ||||
Db      55300 TCAACAGAGTAAATAGACAACCTACAGAATAGAAGAAAATATTTGCAAACTATGCATCCA 55241

Qy         61 ATTAGAGACTTCTATCCAG 79
              |  |  | ||  |||||||
Db      55240 ACAAAGGTCTAATATCCAG 55222


RESULT 11
US-09-288-143-27/c
; Sequence 27, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
;  APPLICANT: Brewer et al.
;  TITLE OF INVENTION: 53 Human Secreted Proteins
;  FILE REFERENCE: PZ018P1
;  CURRENT APPLICATION NUMBER: US/09/288,143
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/US98/21142
;  EARLIER FILING DATE: 1998-10-08
;  EARLIER APPLICATION NUMBER: 60/061,463
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,529
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/071,498
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,527
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,536
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,532
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 219
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
;   LENGTH: 788
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-288-143-27

  Query Match             44.0%;  Score 35.6;  DB 4;  Length 788;
  Best Local Similarity   67.6%;  Pred. No. 0.0045;
  Matches   50;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          6 AAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGATTAG 65
              | |||||||||   ||||  |||||      |||| | || ||  ||||||||  || ||
Db        107 AGAGTGAAAAGGTAACCTACAGAATGGGAGAAAATATTTGCAAACCATGTATCTAATAAG 48

Qy         66 AGACTTCTATCCAG 79
              ||  |  |||||||
Db         47 AGGTTAATATCCAG 34


RESULT 12
US-09-539-333D-211
; Sequence 211, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Bihain, Bernard
;  APPLICANT:  Essioux, Laurent
;  TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
;  FILE REFERENCE: GENSET.047AUS
;  CURRENT APPLICATION NUMBER: US/09/539,333D
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION NUMBER: US 60/126,903
;  PRIOR FILING DATE: 1999-03-30
;  PRIOR APPLICATION NUMBER: US 60/131,971
;  PRIOR FILING DATE: 1999-04-30
;  PRIOR APPLICATION NUMBER: US 60/132,065
;  PRIOR FILING DATE: 1999-04-30
;  PRIOR APPLICATION NUMBER: US 60/143,928
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 60/145,915
;  PRIOR FILING DATE: 1999-07-27
;  PRIOR APPLICATION NUMBER: US 60/146,453
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: US 60/146,452
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: US 60/162,288
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: US 09/416,384
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 231
;  SOFTWARE: Patent.pm
; SEQ ID NO 211
;   LENGTH: 3001
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1501
;   OTHER INFORMATION: 99-26781-25 : polymorphic base G or T
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1482..1500
;   OTHER INFORMATION: 99-26781-25.mis1
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1502..1521
;   OTHER INFORMATION: 99-26781-25.mis2,  complement
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1477..1497
;   OTHER INFORMATION: upstream amplification primer
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1905..1925
;   OTHER INFORMATION: downstream amplification primer, complement
;   FEATURE:
;   NAME/KEY: misc_binding
;   LOCATION: 1489..1513
;   OTHER INFORMATION: 99-26781-25  probe
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 21,274..275
;   OTHER INFORMATION: n=a, g, c or t
US-09-539-333D-211

  Query Match             43.0%;  Score 34.8;  DB 4;  Length 3001;
  Best Local Similarity   65.4%;  Pred. No. 0.012;
  Matches   51;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              | |  ||||| |  | ||||||| ||||   |   ||||| | || ||   ||||||| |
```

```
Db       1313 TGAGCAAAGTAAGCAGACAACCCACAGAGTGAGAGAAAATATTTGCAAACTATGTATCTG 1372

Qy         61 ATTAGAGACTTCTATCCA 78
              |  |  ||||  ||||||
Db       1373 ACAAAGGACTAATATCCA 1390


RESULT 13
US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
;  APPLICANT: GAN, Weiniu
;  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001147
;  CURRENT APPLICATION NUMBER: US/09/816,095
;  CURRENT FILING DATE:  2001-03-26
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 99916
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(99916)
;   OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

  Query Match             42.7%;  Score 34.6;  DB 4;  Length 99916;
  Best Local Similarity   64.2%;  Pred. No. 0.032;
  Matches   52;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||| ||| || | ||||||   ||||      ||| | | |||||   ||| ||| |
Db      47456 TCAAGAGAGTAAACAGACAACCTATAGAATGGGAGAAAGTATTTGTAAAGTATGCATCTG 47515

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              |  |  | ||  |||||| |
Db      47516 ACAAAGGTCTAATATCCAGAA 47536


RESULT 14
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001067
;  CURRENT APPLICATION NUMBER: US/09/751,389
;  CURRENT FILING DATE:  2001-01-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 786431
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(786431)
;   OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

  Query Match             42.7%;  Score 34.6;  DB 4;  Length 786431;
  Best Local Similarity   64.2%;  Pred. No. 0.053;
  Matches   52;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||| | ||||||| | |||      |||||  | ||  ||   |  |||| |
Db     367146 TCAACAAAGTGAATAGACAACCCACTGAATGGGAGAAAATATTTGCAAACTACCTATCTG 367087

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              |  ||  | |  || |||| |
Db     367086 ACAAGGAATTGATACCCAGAA 367066


RESULT 15
US-09-621-976-1442/c
; Sequence 1442, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 1442
;   LENGTH: 523
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 188..472
;   NAME/KEY: sig_peptide
;   LOCATION: 188..295
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 7.80000019073486
;   OTHER INFORMATION: seq CFVLLFFLRWSLT/XC
;   NAME/KEY: misc_feature
;   LOCATION: 20,444
;   OTHER INFORMATION: n=a, g, c or t
US-09-621-976-1442

  Query Match             41.7%;  Score 33.8;  DB 4;  Length 523;
  Best Local Similarity   61.7%;  Pred. No. 0.016;
  Matches   50;  Conservative    2;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              | ||:| | | | || |||| :| || |       |||||  | ||| |||||| |||| |
Db        137 TTAAKAKAATAAKAAKACAAGSCACACATTAGGAKAAAATATTTGTGAGTCATCTATCTG 78

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  |||||   | || |
Db         77 ATAAGSTACTTCCTCCTAGAA 57


Search completed: May  7, 2004, 15:44:43
Job time : 43.6738 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 11:56:28 ; Search time 165.476 Seconds
(without alignments)
2079.475 Million cell updates/sec

Title: US-10-071-411A-1_COPY_500_580
Perfect score: 81
Sequence: 1 tcaagaaagtgaaaacacaa.....ttagagacttctatccagga 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747718

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	60.4	74.6	13249	6	ABL32117	Abl32117 Human imm
c	2	60.4	74.6	13249	6	ABK31177	Abk31177 Signal tr
c	3	60.4	74.6	13249	6	ABL70132	Abl70132 Chemicall
	4	58.6	72.3	13249	6	ABL32116	Abl32116 Human imm
	5	58.6	72.3	13249	6	ABK31176	Abk31176 Signal tr
	6	58.6	72.3	13249	6	ABL70131	Abl70131 Chemicall
c	7	50.2	62.0	72332	8	ADA02552	Ada02552 Human WNT
c	8	50.2	62.0	72332	9	ADB72290	Adb72290 Human WNT
c	9	47.4	58.5	16532	4	AAK77730	Aak77730 Human imm
c	10	47.4	58.5	16535	4	AAK77731	Aak77731 Human imm
c	11	47.4	58.5	128034	9	ADE43582	Ade43582 Polymorph
c	12	47.4	58.5	128034	9	ADE43581	Ade43581 Human IDE
	13	47.4	58.5	202100	9	ADE43315	Ade43315 Human IDE
	14	47	58.0	752	4	AAH04914	Aah04914 Human cDN
	15	47	58.0	1794	4	AAH17530	Aah17530 Human cDN
c	16	46.4	57.3	110000	9	ADE11169_2	Continuation (3 of
c	17	45.8	56.5	432	6	ABK53951	Abk53951 Human hea
c	18	45.8	56.5	478	6	ABK54007	Abk54007 Human hea
c	19	45.8	56.5	523	6	ABK54004	Abk54004 Human hea
c	20	45.8	56.5	524	6	ABK54001	Abk54001 Human hea
c	21	45.8	56.5	649	2	AAZ24559	Aaz24559 Human lun
c	22	45.8	56.5	649	3	AAC65798	Aac65798 Human lun
c	23	45.8	56.5	649	6	ABL49017	Abl49017 Human lun
c	24	45.8	56.5	649	6	ABQ92203	Abq92203 Human lun
c	25	45.8	56.5	649	8	ADA28618	Ada28618 Human lun
c	26	45.8	56.5	649	9	ADE53578	Ade53578 Human lun
c	27	45.8	56.5	1419	4	AAF33258	Aaf33258 Human sec
c	28	45.8	56.5	2625	3	AAC60048	Aac60048 Human sec
c	29	45.8	56.5	2625	7	ADA97959	Ada97959 Human sec
c	30	45.8	56.5	2625	7	ADA43865	Ada43865 Human sec
c	31	45.8	56.5	2625	9	ADC20114	Adc20114 Human sec
	32	45.8	56.5	110000	7	ACF42745_0	Acf42745 Human ALM
c	33	45.4	56.0	20247	4	AAL36315	Aal36315 Human mus
c	34	45.4	56.0	20247	7	ABX59303	Abx59303 cDNA enco
c	35	45	55.6	314	6	ABK53813	Abk53813 Human hea
c	36	44.4	54.8	240000	7	ACD13446	Acd13446 Human DNA
	37	44.2	54.6	197997	7	AAL54074	Aal54074 Human tra
c	38	44.2	54.6	201143	6	ABK83568	Abk83568 Human DNA
c	39	43.8	54.1	486	6	ABK54006	Abk54006 Human hea
	40	42.8	52.8	348	6	ABN22400	Abn22400 Human ORF
	41	42.8	52.8	10093	4	AAK68011	Aak68011 Human imm
	42	42.8	52.8	10093	4	AAL03622	Aal03622 Human rep
	43	42.8	52.8	10093	4	ABA07820	Aba07820 Human ova
c	44	42.6	52.6	384	4	AAI81654	Aai81654 Human pol
	45	42.6	52.6	401	4	AAK95385	Aak95385 Human neu

ALIGNMENTS

RESULT 1
ABL32117/c
ID ABL32117 standard; DNA; 13249 BP.
XX
AC ABL32117;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 90.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 90; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

Query Match 74.6%; Score 60.4; DB 6; Length 13249;
Best Local Similarity 85.9%; Pred. No. 8.6e-10;
Matches 67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
Db 9541 TCAAAAAAATAAAAACACAACCCGCAAAAACAATAAAAATATCTATAAATCATATATCCG 9482

Qy 61 ATTAGAGACTTCTATCCA 78
Db 9481 ATTAAAAACTTCTATCCA 9464

RESULT 2
ABK31177/c
ID ABK31177 standard; DNA; 13249 BP.
XX
AC ABK31177;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #10.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007472.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction.
XX
PS Claim 1; SEQ ID NO 20; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC the cytosine methylation state (CpG islands) of these genes, and a method
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC genes associated with signal transduction. The genomic DNA can be
CC obtained from cells or cellular components which contain DNA, e.g. cell
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC brain, heart, prostate, lung, breast or liver, histologic object slides,
CC and all their possible combinations. The sequences of the invention are
CC useful for the diagnosis and therapy of diseases associated with signal
CC transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC chemically pretreated genomic DNA sequences of different genes associated
CC with signal transduction, or their complementary sequences. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

Query Match 74.6%; Score 60.4; DB 6; Length 13249;
Best Local Similarity 85.9%; Pred. No. 8.6e-10;
Matches 67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
Db 9541 TCAAAAAAATAAAAACACAACCCGCAAAAACAATAAAAATATCTATAAATCATATATCCG 9482

Qy 61 ATTAGAGACTTCTATCCA 78
Db 9481 ATTAAAAACTTCTATCCA 9464

RESULT 3
ABL70132/c
ID ABL70132 standard; DNA; 13249 BP.
XX
AC ABL70132;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence complementary to#11.
XX
KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007471.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 22; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX

```
SQ   Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

  Query Match             74.6%;  Score 60.4;  DB 6;  Length 13249;
  Best Local Similarity   85.9%;  Pred. No. 8.6e-10;
  Matches   67;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| ||| | ||||||||||||||| || |||||||||| ||| ||| |||| ||||||
Db       9541 TCAAAAAAATAAAAACACAACCCGCAAAAACAATAAAAATATCTATAAATCATATATCCG 9482

Qy         61 ATTAGAGACTTCTATCCA 78
              |||| | ||||||||||
Db       9481 ATTAAAAACTTCTATCCA 9464


RESULT 4
ABL32116
ID   ABL32116 standard; DNA; 13249 BP.
XX
AC   ABL32116;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 89.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 89; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 58.6;  DB 6;  Length 13249;
  Best Local Similarity   82.7%;  Pred. No. 3.3e-09;
  Matches   67;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              | |||||||||||||| | ||  || ||||| |||||||||| ||||||| |||||| ||
Db       3709 TTAAGAAAGTGAAAATATAATTCGTAGAAGTAATAAAAATGTTTGTAAGTTATGTATTCG 3768

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              |||||||| || |||  ||||
Db       3769 ATTAGAGATTTTTATTTAGGA 3789


RESULT 5
ABK31176
ID   ABK31176 standard; DNA; 13249 BP.
XX
AC   ABK31176;
XX
DT   23-APR-2002  (first entry)
XX
DE   Signal transduction associated gene modified DNA #10.
XX
KW   Human; signal transduction associated gene; cytosine methylation state;
KW   CpG island; signal transduction associated disease; solid tumour; cancer;
KW   antitumour; cytostatic; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007472.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction.
XX
PS   Claim 1; SEQ ID NO 19; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC   Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC   the cytosine methylation state (CpG islands) of these genes, and a method
CC   for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC   genes associated with signal transduction. The genomic DNA can be
CC   obtained from cells or cellular components which contain DNA, e.g. cell
CC   lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC   tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC   brain, heart, prostate, lung, breast or liver, histologic object slides,
CC   and all their possible combinations. The sequences of the invention are
CC   useful for the diagnosis and therapy of diseases associated with signal
CC   transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC   chemically pretreated genomic DNA sequences of different genes associated
CC   with signal transduction, or their complementary sequences. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 58.6;  DB 6;  Length 13249;
  Best Local Similarity   82.7%;  Pred. No. 3.3e-09;
  Matches   67;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
```

```
Db       3709 TTAAGAAAGTGAAAATATAATTCGTAGAAGTAATAAAAATGTTTGTAAGTTATGTATTCG 3768

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              ||||||||  ||  |||  ||||
Db       3769 ATTAGAGATTTTTATTTAGGA 3789
```

RESULT 6
ABL70131
ID ABL70131 standard; DNA; 13249 BP.
XX
AC ABL70131;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence#11.
XX
KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007471.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 21; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

```
  Query Match             72.3%;  Score 58.6;  DB 6;  Length 13249;
  Best Local Similarity   82.7%;  Pred. No. 3.3e-09;
  Matches   67;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||||||||||||| | ||  || ||||| |||||||||||| ||||||| |||||| ||
Db       3709 TTAAGAAAGTGAAAATATAATTCGTAGAAGTAATAAAAATGTTTGTAAGTTATGTATTCG 3768

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              ||||||||  ||  |||  ||||
```

```
Db       3769 ATTAGAGATTTTTATTTAGGA 3789
```

RESULT 7
ADA02552/c
ID ADA02552 standard; DNA; 72332 BP.
XX
AC ADA02552;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human WNT3A carcinoma associated gene, SEQ ID NO:1070.
XX
KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1070; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed human CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 72332 BP; 16680 A; 18843 C; 19431 G; 16654 T; 0 U; 724 Other;

```
  Query Match             62.0%;  Score 50.2;  DB 8;  Length 72332;
  Best Local Similarity   77.2%;  Pred. No. 2.5e-06;
  Matches   61;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||| |||||||||| ||||||| |||||       ||||| | || |||||||||||  |
Db      36015 TCAAGCAAGTGAAAAGACAACCCACAGAATGGGGGAAAATATTTGCAAGTCATGTATTGG 35956

Qy         61 ATTAGAGACTTCTATCCAG 79
              || || ||||| ||| |||
```

```
Db      35955 ATAAGGGACTTGTATGCAG 35937


RESULT 8
ADB72290/c
ID   ADB72290 standard; DNA; 72332 BP.
XX
AC   ADB72290;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human WNT3A gene.
XX
KW   human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
PR   08-NOV-2001; 2001US-00052482.
PR   30-NOV-2001; 2001US-00997722.
PR   20-DEC-2001; 2001US-00034650.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-239337/23.
XX
PT   New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT   cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS   Claim 1; SEQ ID NO 118; 2304pp; English.
XX
CC   The invention relates to a novel recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the 660 sequences fully defined
CC   in the specification. A polynucleotide of the invention has cytostatic
CC   activity, and may have a use in gene therapy, or in a vaccine. The
CC   recombinant nucleic acids and polypeptides are useful for treating
CC   carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC   sarcomas. The present sequence represents a human gene of the invention.
XX
SQ   Sequence 72332 BP; 16680 A; 18843 C; 19431 G; 16654 T; 0 U; 724 Other;

  Query Match             62.0%;  Score 50.2;  DB 9;  Length 72332;
  Best Local Similarity   77.2%;  Pred. No. 2.5e-06;
  Matches   61;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||| ||||||||| ||||||| |||||      |||||  | ||  ||||||||||  |
Db      36015 TCAAGCAAGTGAAAAGACAACCCACAGAATGGGGGAAAATATTTGCAAGTCATGTATTGG 35956

Qy         61 ATTAGAGACTTCTATCCAG 79
              || || ||||| ||| |||
Db      35955 ATAAGGGACTTGTATGCAG 35937


RESULT 9
AAK77730/c
ID   AAK77730 standard; DNA; 16532 BP.
XX
AC   AAK77730;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32542.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
```

PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 32542; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 16532 BP; 4442 A; 3253 C; 3304 G; 5533 T; 0 U; 0 Other;

Query Match 58.5%; Score 47.4; DB 4; Length 16532;
Best Local Similarity 74.1%; Pred. No. 1.6e-05;
Matches 60; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||| | |||||      |||||  | || || |||||||||  |
Db       4021 TCAAAAAAGTGAAAAGACAACTCACAGAAAGGGAGAAAATATTTGCAAATCATGTATCTG 3962

Qy         61 ATTAGAGACTTCTATCCAGGA 81
                || || ||||| |  |||| |
Db       3961 ATAAGGGACTTGTGCCCAGAA 3941
```

RESULT 10
AAK77731/c
ID AAK77731 standard; DNA; 16535 BP.
XX
AC AAK77731;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32543.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 32543; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)

```
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 16535 BP; 4440 A; 3253 C; 3307 G; 5535 T; 0 U; 0 Other;

  Query Match           58.5%;  Score 47.4;  DB 4;  Length 16535;
  Best Local Similarity  74.1%;  Pred. No. 1.6e-05;
  Matches   60;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              |||| |||||||||| ||||| | |||||      |||||  | || || ||||||||| |
Db       4022 TCAAAAAAGTGAAAAGACAACTCACAGAAAGGGAGAAAATATTTGCAAATCATGTATCTG 3963

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || ||  ||||| |  |||| |
Db       3962 ATAAGGGACTTGTGCCCAGAA 3942


RESULT 11
ADE43582/c
ID   ADE43582 standard; DNA; 128034 BP.
XX
AC   ADE43582;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Polymorphic human IDE genomic sequence, SEQ ID 187.
XX
KW   Neurodegenerative disease; uPA; SNCG; IDE; KNSL1; LIPA; TNFRSF6;
KW   Alzheimer's disease; neuroprotective; nootropic; gene therapy;
KW   Chromosome 10; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    2456
FT                   /*tag=  a
FT                   /note= "There is a variation at this position"
FT   misc_feature    3279
FT                   /*tag=  b
FT                   /note= "There is a variation at this position"
FT   misc_feature    3407
FT                   /*tag=  c
FT                   /note= "There is a variation at this position"
FT   misc_feature    17095
FT                   /*tag=  d
FT                   /note= "There is a variation at this position"
FT   misc_feature    17242
FT                   /*tag=  e
FT                   /note= "There is a variation at this position"
FT   misc_feature    33590
FT                   /*tag=  f
FT                   /note= "There is a variation at this position"
FT   misc_feature    38903
FT                   /*tag=  g
FT                   /note= "There is a variation at this position"
FT   misc_feature    42943
FT                   /*tag=  h
FT                   /note= "There is a variation at this position"
FT   misc_feature    43391
FT                   /*tag=  i
FT                   /note= "There is a variation at this position"
FT   misc_feature    45017
FT                   /*tag=  j
FT                   /note= "There is a variation at this position"
FT   misc_feature    62498
FT                   /*tag=  k
FT                   /note= "There is a variation at this position"
FT   misc_feature    68906
FT                   /*tag=  l
FT                   /note= "There is a variation at this position"
FT   misc_feature    68973
FT                   /*tag=  m
FT                   /note= "There is a variation at this position"
FT   misc_feature    69586
FT                   /*tag=  n
FT                   /note= "There is a variation at this position"
FT   misc_feature    73772
FT                   /*tag=  o
FT                   /note= "There is a variation at this position"
FT   misc_feature    74084
FT                   /*tag=  p
FT                   /note= "There is a variation at this position"
FT   misc_feature    83024
FT                   /*tag=  q
FT                   /note= "There is a variation at this position"
FT   misc_feature    83104
FT                   /*tag=  r
FT                   /note= "There is a variation at this position"
FT   misc_feature    89301
FT                   /*tag=  s
FT                   /note= "There is a variation at this position"
FT   misc_feature    105060
FT                   /*tag=  t
FT                   /note= "There is a variation at this position"
FT   misc_feature    107395
FT                   /*tag=  u
FT                   /note= "There is a variation at this position"
FT   misc_feature    108489
FT                   /*tag=  v
FT                   /note= "There is a variation at this position"
FT   misc_feature    111914
FT                   /*tag=  w
FT                   /note= "There is a variation at this position"
FT   misc_feature    112114
FT                   /*tag=  x
FT                   /note= "There is a variation at this position"
FT   misc_feature    113142
FT                   /*tag=  y
FT                   /note= "There is a variation at this position"
FT   misc_feature    113591
FT                   /*tag=  z
FT                   /note= "There is a variation at this position"
FT   misc_feature    114683
FT                   /*tag=  aa
FT                   /note= "There is a variation at this position"
FT   misc_feature    116662
FT                   /*tag=  ab
FT                   /note= "There is a variation at this position"
FT   misc_feature    117803
FT                   /*tag=  ac
FT                   /note= "There is a variation at this position"
FT   misc_feature    124565
FT                   /*tag=  ad
FT                   /note= "There is a variation at this position"
XX
PN   WO2003054143-A2.
XX
PD   03-JUL-2003.
XX
PF   25-OCT-2002; 2002WO-US034679.
XX
PR   25-OCT-2001; 2001US-0339525P.
PR   08-NOV-2001; 2001US-0336929P.
```

PR 08-NOV-2001; 2001US-0338010P.
PR 09-NOV-2001; 2001US-0338363P.
PR 04-DEC-2001; 2001US-0337052P.
PR 28-MAR-2002; 2002US-0368919P.
XX
PA (NEUR-) NEUROGENETICS INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Becker KD, Velicelebi G, Elliott KJ, Wang X, Tanzi RE, Bertram L;
PI Saunders AJ, Mullin KM, Sampson AJ, Blacker DL;
XX
DR WPI; 2003-559131/52.
XX
PT Determining a predisposition for or the occurrence of neurodegenerative
PT disease, e.g. Alzheimer's disease by detecting in a target nucleic acid
PT the presence or absence of an allelic variant of one or more polymorphic
PT regions.
XX
PS Claim 9; Page 584-page 618; 848pp; English.
XX
CC The present invention relates to a method (M1) for determining a
CC predisposition for or the occurrence of neurodegenerative disease in a
CC subject. The method comprises detecting in a target nucleic acid obtained
CC from the subject the presence or absence of an allelic variant of one or
CC more polymorphic regions of one or more genes selected from uPA
CC (Urokinase plasminogen activator), SNCG (gamma-synuclein), IDE (insulin-
CC degrading enzyme), KNSL1 (Kinesin-like protein 1), LIPA (lysosomal acid
CC lypase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the
CC presence of at least one of the allelic variant of one or more
CC polymorphic regions is indicative of a predisposition for or the
CC occurrence of neurodegenerative disease. The genes are all located on
CC chromosome 10. M1 is useful for determining a predisposition for or the
CC occurrence of, and for treating neurodegenerative disease, particularly
CC Alzheimer's disease.
XX
SQ Sequence 128034 BP; 34726 A; 25977 C; 26400 G; 40799 T; 0 U; 132 Other;

Query Match 58.5%; Score 47.4; DB 9; Length 128034;
Best Local Similarity 74.1%; Pred. No. 2.2e-05;
Matches 60; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
Db 18972 TCAAGAAAGTGAAAAGACAACCTATAGAATGGCATAAAATATTTGTAAATCATATATCTG 18913

Qy 61 ATTAGAGACTTCTATCCAGGA 81
Db 18912 ATAAGGGACTTGTATAGAGAA 18892

RESULT 12
ADE43581/c
ID ADE43581 standard; DNA; 128034 BP.
XX
AC ADE43581;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human IDE genomic sequence, SEQ ID 186.
XX
KW Neurodegenerative disease; uPA; SNCG; IDE; KNSL1; LIPA; TNFRSF6;
KW Alzheimer's disease; neuroprotective; nootropic; gene therapy;
KW Chromosome 10; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003054143-A2.
XX
PD 03-JUL-2003.
XX
PF 25-OCT-2002; 2002WO-US034679.
XX
PR 25-OCT-2001; 2001US-0339525P.

PR 08-NOV-2001; 2001US-0336929P.
PR 08-NOV-2001; 2001US-0338010P.
PR 09-NOV-2001; 2001US-0338363P.
PR 04-DEC-2001; 2001US-0337052P.
PR 28-MAR-2002; 2002US-0368919P.
XX
PA (NEUR-) NEUROGENETICS INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Becker KD, Velicelebi G, Elliott KJ, Wang X, Tanzi RE, Bertram L;
PI Saunders AJ, Mullin KM, Sampson AJ, Blacker DL;
XX
DR WPI; 2003-559131/52.
XX
PT Determining a predisposition for or the occurrence of neurodegenerative
PT disease, e.g. Alzheimer's disease by detecting in a target nucleic acid
PT the presence or absence of an allelic variant of one or more polymorphic
PT regions.
XX
PS Claim 22; Page 549-584; 848pp; English.
XX
CC The present invention relates to a method (M1) for determining a
CC predisposition for or the occurrence of neurodegenerative disease in a
CC subject. The method comprises detecting in a target nucleic acid obtained
CC from the subject the presence or absence of an allelic variant of one or
CC more polymorphic regions of one or more genes selected from uPA
CC (Urokinase plasminogen activator), SNCG (gamma-synuclein), IDE (insulin-
CC degrading enzyme), KNSL1 (Kinesin-like protein 1), LIPA (lysosomal acid
CC lypase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the
CC presence of at least one of the allelic variant of one or more
CC polymorphic regions is indicative of a predisposition for or the
CC occurrence of neurodegenerative disease. The genes are all located on
CC chromosome 10. M1 is useful for determining a predisposition for or the
CC occurrence of, and for treating neurodegenerative disease, particularly
CC Alzheimer's disease.
XX
SQ Sequence 128034 BP; 34731 A; 25985 C; 26409 G; 40808 T; 0 U; 101 Other;

Query Match 58.5%; Score 47.4; DB 9; Length 128034;
Best Local Similarity 74.1%; Pred. No. 2.2e-05;
Matches 60; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
Db 18972 TCAAGAAAGTGAAAAGACAACCTATAGAATGGCATAAAATATTTGTAAATCATATATCTG 18913

Qy 61 ATTAGAGACTTCTATCCAGGA 81
Db 18912 ATAAGGGACTTGTATAGAGAA 18892

RESULT 13
ADE43315
ID ADE43315 standard; DNA; 202100 BP.
XX
AC ADE43315;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human IDE/ KNSL1 genomic sequence, SEQ ID 484.
XX
KW Neurodegenerative disease; uPA; SNCG; IDE; KNSL1; LIPA; TNFRSF6;
KW Alzheimer's disease; neuroprotective; nootropic; gene therapy;
KW Chromosome 10; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003054143-A2.
XX
PD 03-JUL-2003.
XX
PF 25-OCT-2002; 2002WO-US034679.
XX

PR 25-OCT-2001; 2001US-0339525P.
PR 08-NOV-2001; 2001US-0336929P.
PR 08-NOV-2001; 2001US-0338010P.
PR 09-NOV-2001; 2001US-0338363P.
PR 04-DEC-2001; 2001US-0337052P.
PR 28-MAR-2002; 2002US-0368919P.
XX
PA (NEUR-) NEUROGENETICS INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Becker KD, Velicelebi G, Elliott KJ, Wang X, Tanzi RE, Bertram L;
PI Saunders AJ, Mullin KM, Sampson AJ, Blacker DL;
XX
DR WPI; 2003-559131/52.
XX
PT Determining a predisposition for or the occurrence of neurodegenerative
PT disease, e.g. Alzheimer's disease by detecting in a target nucleic acid
PT the presence or absence of an allelic variant of one or more polymorphic
PT regions.
XX
PS Claim 9; Page 769-823; 848pp; English.
XX
CC The present invention relates to a method (M1) for determining a
CC predisposition for or the occurrence of neurodegenerative disease in a
CC subject. The method comprises detecting in a target nucleic acid obtained
CC from the subject the presence or absence of an allelic variant of one or
CC more polymorphic regions of one or more genes selected from uPA
CC (Urokinase plasminogen activator), SNCG (gamma-synuclein), IDE (insulin-
CC degrading enzyme), KNSL1 (Kinesin-like protein 1), LIPA (lysosomal acid
CC lypase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the
CC presence of at least one of the allelic variant of one or more
CC polymorphic regions is indicative of a predisposition for or the
CC occurrence of neurodegenerative disease. The genes are all located on
CC chromosome 10. M1 is useful for determining a predisposition for or the
CC occurrence of, and for treating neurodegenerative disease, particularly
CC Alzheimer's disease.
XX
SQ Sequence 202100 BP; 60747 A; 41352 C; 41113 G; 58888 T; 0 U; 0 Other;

Query Match 58.5%; Score 47.4; DB 9; Length 202100;
Best Local Similarity 74.1%; Pred. No. 2.4e-05;
Matches 60; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy          1 TCAAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCG 60
              ||||||||||||||| ||||||   ||||      ||||| | ||||| |||| |||| |
Db     104723 TCAAGAAAGTGAAAAGACAACCTATAGAATGGCATAAAATATTTGTAAATCATATATCTG 104782

Qy         61 ATTAGAGACTTCTATCCAGGA 81
              || || ||||| |||  || |
Db     104783 ATAAGGGACTTGTATAGAGAA 104803
```

RESULT 14
AAH04914
ID AAH04914 standard; cDNA; 752 BP.
XX
AC AAH04914;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1749.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 1749; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 752 BP; 289 A; 140 C; 137 G; 183 T; 0 U; 3 Other;

Query Match 58.0%; Score 47; DB 4; Length 752;
Best Local Similarity 74.7%; Pred. No. 1.2e-05;
Matches 59; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              || |||||||||| |||||||  ||||  |  |||||| | |  ||| ||||||||| |||
Db        620 AAAAAAGTGAAAATACAACCCATAGAAAGATAAAAAATATTTTCAAGCCATGTATCTGAT 679

Qy         63 TAGAGACTTCTATCCAGGA 81
              || | ||  |||||||  |
Db        680 AAGGGTCTAGTATCCAGAA 698
```

RESULT 15
AAH17530
ID AAH17530 standard; cDNA; 1794 BP.
XX
AC AAH17530;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:17007.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX

PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 17007; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1794 BP; 665 A; 328 C; 368 G; 433 T; 0 U; 0 Other;

Query Match 58.0%; Score 47; DB 4; Length 1794;
Best Local Similarity 74.7%; Pred. No. 1.4e-05;
Matches 59; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          3 AAGAAAGTGAAAACACAACCCGCAGAAGCAATAAAAATGTCTGTAAGTCATGTATCCGAT 62
              || |||||||||| |||||||  ||||  |  |||||| | |  ||| |||||||| |||
Db        620 AAAAAAGTGAAAATACAACCCATAGAAAGATAAAAAATATTTTCAAGCCATGTATCTGAT 679

Qy         63 TAGAGACTTCTATCCAGGA 81
                || | ||  ||||||| |
Db        680 AAGGGTCTAGTATCCAGAA 698
```

Search completed: May 7, 2004, 13:50:26
Job time : 168.476 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:34:53 ; Search time 934.416 Seconds
(without alignments)
1629.864 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	34	66.7	313	28	AQ164003	AQ164003 HS_2270_B
c	2	30.6	60.0	416	28	AQ442281	AQ442281 HS_5137_A
	3	30.6	60.0	723	29	CE266277	CE266277 tigr-gss-
	4	30.2	59.2	815	29	CC535951	CC535951 CH240_414
c	5	29.8	58.4	491	28	AQ047026	AQ047026 RPCI11-35
c	6	29	56.9	382	28	AQ349913	AQ349913 RPCI11-12
c	7	29	56.9	809	29	CC489556	CC489556 CH240_322
	8	28.8	56.5	642	29	CE520001	CE520001 tigr-gss-
	9	28.6	56.1	479	29	CE454299	CE454299 tigr-gss-
c	10	28.2	55.3	265	28	AQ106326	AQ106326 HS_3070_A
	11	27.8	54.5	669	29	CE593552	CE593552 tigr-gss-
c	12	27.6	54.1	841	28	BZ387096	BZ387096 EINCM29TF
c	13	27.4	53.7	646	29	AG048782	AG048782 Pan trogl
c	14	27.4	53.7	700	28	AQ386453	AQ386453 RPCI11-15
	15	27.2	53.3	449	28	AQ887236	AQ887236 HS_5552_B
c	16	27.2	53.3	455	28	AZ043603	AZ043603 RPCI-23-3
c	17	27.2	53.3	738	29	CC471346	CC471346 CH240_294
c	18	27	52.9	312	9	AI239979	AI239979 qh44d09.x
	19	27	52.9	351	13	BX099859	BX099859 BX099859
c	20	27	52.9	355	13	BU763852	BU763852 sas48h09.
c	21	27	52.9	362	10	AW183073	AW183073 xj66e05.x
c	22	27	52.9	363	10	BE041580	BE041580 ho51g12.x
c	23	27	52.9	410	10	AW628725	AW628725 hi43b04.x
c	24	27	52.9	419	28	AQ817217	AQ817217 HS_5274_B
	25	27	52.9	521	29	CE297174	CE297174 tigr-gss-
c	26	27	52.9	557	29	CE104592	CE104592 tigr-gss-
c	27	27	52.9	672	29	CE299189	CE299189 tigr-gss-
c	28	27	52.9	678	9	AV981607	AV981607 AV981607
	29	27	52.9	693	9	AV869864	AV869864 AV869864
	30	26.8	52.5	759	14	CK028932	CK028932 AGENCOURT
c	31	26.8	52.5	846	29	CG960324	CG960324 MBENC49TR
c	32	26.6	52.2	412	10	AW283536	AW283536 LG1_285_G
	33	26.6	52.2	550	28	BZ339389	BZ339389 ic31h04.g
c	34	26.6	52.2	596	14	CD210389	CD210389 HS1_51_C0
c	35	26.6	52.2	619	14	CB084591	CB084591 hq18b10.b
	36	26.6	52.2	737	28	BH203848	BH203848 Sm1-44K7.
c	37	26.6	52.2	833	14	CD757791	CD757791 AGENCOURT
	38	26.2	51.4	203	10	BB277304	BB277304 BB277304
	39	26.2	51.4	356	10	BE801621	BE801621 sr16g05.y
c	40	26.2	51.4	468	13	BY512389	BY512389 BY512389
	41	26.2	51.4	524	29	CC676917	CC676917 OGUDB34TV
	42	26.2	51.4	550	28	AQ534651	AQ534651 RPCI-11-3
	43	26.2	51.4	633	29	CE438117	CE438117 tigr-gss-
	44	26.2	51.4	688	28	AQ350492	AQ350492 RPCI11-12
	45	26.2	51.4	702	29	BX198925	BX198925 Danio rer

ALIGNMENTS

RESULT 1
AQ164003
LOCUS AQ164003 313 bp DNA linear GSS 16-OCT-1998
DEFINITION HS_2270_B1_E07_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2270 Col=13 Row=J, genomic survey sequence.
ACCESSION AQ164003
VERSION AQ164003.1 GI:3562198
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 313)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618

```
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2270  row: J  column: 13
            Class: BAC ends
            High quality sequence stop: 313.
FEATURES              Location/Qualifiers
     source           1. .313
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=2270 Col=13 Row=J"
                      /sex="male"
                      /clone_lib="CIT Approved Human Genomic Sperm Library D"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
ORIGIN

  Query Match             66.7%;  Score 34;  DB 28;  Length 313;
  Best Local Similarity   80.0%;  Pred. No. 2.8;
  Matches   40;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||||| | ||||||||||||||| ||||| |||| |||| |||||   ||
Db        164 ACAAAGATAAATTGGACTTAAAATTAAAAACTTGTGTGTGTCAAAGGACA 213


RESULT 2
AQ442281/c
LOCUS       AQ442281                416 bp    DNA     linear   GSS 31-MAR-1999
DEFINITION  HS_5137_A1_G07_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=713 Col=13 Row=M, genomic survey sequence.
ACCESSION   AQ442281
VERSION     AQ442281.1  GI:4553620
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 713  row: M  column: 13
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 416.
FEATURES              Location/Qualifiers
     source           1. .416
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="Plate=713 Col=13 Row=M"
                      /sex="male"
                      /clone_lib="RPCI-11 Human Male BAC Library"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      Male blood DNA was isolated from one randomly chosen donor
                      and partially digested with a combination of EcoRI and
                      EcoRI Methylase. Size selected DNA was cloned into the
                      pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             60.0%;  Score 30.6;  DB 28;  Length 416;
  Best Local Similarity   80.0%;  Pred. No. 28;
  Matches   36;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | | | | ||||||||||||||  ||||| || |||||||| |||||
Db        388 AAATAGATAAATTGGACTTAATTTTAAAAACGTTTGTGCCTCAAA 344


RESULT 3
CE266277
LOCUS       CE266277                723 bp    DNA     linear   GSS 26-SEP-2003
DEFINITION  tigr-gss-dog-17000337092175 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE266277
VERSION     CE266277.1  GI:35989516
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES              Location/Qualifiers
     source           1. .723
                      /organism="Canis familiaris"
                      /mol_type="genomic DNA"
                      /strain="Standard Poodle"
                      /db_xref="taxon:9615"
                      /clone_lib="Dog Library"
                      /note="Site_1: BstXI; Libraries were prepared from
                      peripheral blood"
ORIGIN

  Query Match             60.0%;  Score 30.6;  DB 29;  Length 723;
  Best Local Similarity   80.0%;  Pred. No. 26;
  Matches   36;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | | ||| | | | |||||||| ||||| |||||||||| |||||
Db        424 AGACAAAAATAGTTGACTTAAAATTAAAAACTTTTGTGCATCAAA 468


RESULT 4
CC535951
LOCUS       CC535951                815 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_414P12.T7 CHORI-240 Bos taurus genomic clone CH240_414P12,
            genomic survey sequence.
ACCESSION   CC535951
VERSION     CC535951.1  GI:31854239
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
```

```
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 815)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_414P12.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 414  row: P  column: 12
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .815
                      /organism="Bos taurus"
                      /mol_type="genomic DNA"
                      /strain="breed: Hereford"
                      /db_xref="taxon:9913"
                      /clone="CH240_414P12"
                      /sex="Male"
                      /cell_type="Blood"
                      /clone_lib="CHORI-240"
                      /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                      Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                      library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             59.2%;  Score 30.2;  DB 29;  Length 815;
  Best Local Similarity   74.5%;  Pred. No. 33;
  Matches   38;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              | |||||  |   |    |||||||||||| |||||||| ||||||| | |||
Db        274 AAAAAAACAGGATTTCATTAAAGTTAAAAACTTTTGTACTTCAAAGAACAT 324


RESULT 5
AQ047026/c
LOCUS       AQ047026                491 bp    DNA     linear   GSS 14-APR-1999
DEFINITION  RPCI11-35C18.TK RPCI-11 Homo sapiens genomic clone RPCI-11-35C18,
            genomic survey sequence.
ACCESSION   AQ047026
VERSION     AQ047026.1  GI:3315953
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-35C18.TJ
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .491
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="GDB:7513121"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-35C18"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /clone_lib="RPCI-11"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             58.4%;  Score 29.8;  DB 28;  Length 491;
  Best Local Similarity   82.9%;  Pred. No. 47;
  Matches   34;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              |||| | | || |||||| ||||||| | ||||||||||||
Db        351 AAAACAGACTGAACTTAATGTTAAAAAATTTTGTGCTTCAA 311


RESULT 6
AQ349913/c
LOCUS       AQ349913                382 bp    DNA     linear   GSS 07-MAY-1999
DEFINITION  RPCI11-120E14.TV RPCI-11 Homo sapiens genomic clone RPCI-11-120E14,
            genomic survey sequence.
ACCESSION   AQ349913
VERSION     AQ349913.1  GI:4177248
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-120E14.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
```

```
     source          1. .382
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7545805"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-120E14"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             56.9%;  Score 29;  DB 28;  Length 382;
  Best Local Similarity   77.8%;  Pred. No. 84;
  Matches   35;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | | | | ||||||||||||||  ||||| || | ||||| |||||
Db        347 AAATAGATAAATTGGACTTAATTTTAAAAACGTGTGTGCCTCAAA 303


RESULT 7
CC489556/c
LOCUS       CC489556                809 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_322D15.T7 CHORI-240 Bos taurus genomic clone CH240_322D15,
            genomic survey sequence.
ACCESSION   CC489556
VERSION     CC489556.1  GI:31800470
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 809)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_322D15.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 322  row: D  column: 15
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .809
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_322D15"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             56.9%;  Score 29;  DB 29;  Length 809;
  Best Local Similarity   77.8%;  Pred. No. 75;
  Matches   35;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | | | | ||||||||||||  | ||| | ||||||||||||| |
Db        348 AGATATATAAACTGGACTTAAAATGAAAAATTTTTGTGCTTCATA 304


RESULT 8
CE520001
LOCUS       CE520001                642 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000365620224 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE520001
VERSION     CE520001.1  GI:36836782
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             56.5%;  Score 28.8;  DB 29;  Length 642;
  Best Local Similarity   75.0%;  Pred. No. 89;
  Matches   36;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              | ||||| |||  | | | |||||||  || |||||||  ||||||  ||
Db         43 AAAAAAAAAAAAGAGAAATAAAAGTTACATTCTTTTGTTGTTCAAAAAT 90


RESULT 9
CE454299
LOCUS       CE454299                479 bp    DNA     linear   GSS 27-SEP-2003
DEFINITION  tigr-gss-dog-17000319453122 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE454299
VERSION     CE454299.1  GI:36747668
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             56.1%;  Score 28.6;  DB 29;  Length 479;
  Best Local Similarity   79.1%;  Pred. No. 1.1e+02;
  Matches   34;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              ||| | |||| |||||||| ||||||   ||||||| ||| ||
Db        152 AAAGAAAAATAGGACTTAACGTTAAAATGTTTTGTGTTTCCAA 194


RESULT 10
AQ106326/c
LOCUS       AQ106326                265 bp    DNA     linear   GSS 28-AUG-1998
DEFINITION  HS_3070_A2_E04_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3070 Col=8 Row=I, genomic survey
            sequence.
ACCESSION   AQ106326
VERSION     AQ106326.1  GI:3481682
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3070  row: I  column: 8
            Class: BAC ends
            High quality sequence stop: 265.
FEATURES             Location/Qualifiers
     source          1. .265
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3070 Col=8 Row=I"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             55.3%;  Score 28.2;  DB 28;  Length 265;
  Best Local Similarity   80.5%;  Pred. No. 1.5e+02;
  Matches   33;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              |||  | | || |||||| |||||| | ||||||||||||
Db        219 AAACCAGACTGAACTTAATGTTAAAAAATTTTGTGCTTCAA 179


RESULT 11
CE593552
LOCUS       CE593552                669 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000366522943 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE593552
VERSION     CE593552.1  GI:36910333
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             54.5%;  Score 27.8;  DB 29;  Length 669;
  Best Local Similarity   82.1%;  Pred. No. 1.8e+02;
  Matches   32;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGC 39
              ||| ||  |||||||||| ||| ||||| | |||||||||
Db        124 ACATAAGTAAATTGGACTAAAAATTAAAAATTTTTGTGC 162


RESULT 12
BZ387096/c
LOCUS       BZ387096                841 bp    DNA     linear   GSS 30-APR-2003
DEFINITION  EINCM29TF EI_10_12_KB Entamoeba invadens genomic clone EINCM29,
            genomic survey sequence.
ACCESSION   BZ387096
VERSION     BZ387096.1  GI:30233424
```

```
KEYWORDS    GSS.
SOURCE      Entamoeba invadens
  ORGANISM  Entamoeba invadens
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   Wang,Z., Samuelson,J., Clark,C.G., Eichinger,D., Paul,J., van
            Dellen,K., Hall,N., Anderson,I. and Loftus,B.
  TITLE     Gene discovery in the Entamoeba invadens genome
  JOURNAL   Mol. Biochem. Parasitol. 129 (1), 23-31 (2003)
  MEDLINE   22684048
   PUBMED   12798503
COMMENT     Other_GSSs: EINCM29TR
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            DNA was provided by Daniel Eichinger
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .841
                     /organism="Entamoeba invadens"
                     /mol_type="genomic DNA"
                     /strain="IP-1"
                     /db_xref="taxon:33085"
                     /clone="EINCM29"
                     /clone_lib="EI_10_12_KB"
                     /note="Vector: pHOS2; Site_1: BstXI; Total genomic DNA was
                     isolated from early log phase trophozoites of E. invadens
                     IP-1 using a Qiagen plant DNA extraction kit. A shotgun
                     medium-size plasmid library (average insert size of 10 -
                     12 kb) was generated by random mechanical shearing of E.
                     invadens genomic DNA, repairing the ends of DNA fragments
                     with T4 Polymerase, adding BstXI adaptors and ligating
                     into the BstXI site of a pUC-derived vector pHOS2."
ORIGIN

  Query Match             54.1%;  Score 27.6;  DB 28;  Length 841;
  Best Local Similarity   72.0%;  Pred. No. 2e+02;
  Matches   36;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              |||||||||| |   | ||||| ||||||| ||||| | |   || || |
Db        458 ACAAAAAGAATTAAAAATTAAATTTAAATATTTTTGAGGTGTGAATATTA 409


RESULT 13
AG048782/c
LOCUS       AG048782                646 bp    DNA     linear   GSS 02-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-028L09.R, genomic survey sequence.
ACCESSION   AG048782
VERSION     AG048782.1  GI:16585674
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 646)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .646
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-028L09.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             53.7%;  Score 27.4;  DB 29;  Length 646;
  Best Local Similarity   75.6%;  Pred. No. 2.3e+02;
  Matches   34;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | |||| | ||||  | | |||||||||| |||| ||| | |||||
Db        385 AAAAAATGGAATTACATTAAAAGTTAAAAACTTCTGTACATCAAA 341


RESULT 14
AQ386453/c
LOCUS       AQ386453                700 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-154D12.TV RPCI-11 Homo sapiens genomic clone RPCI-11-154D12,
            genomic survey sequence.
ACCESSION   AQ386453
VERSION     AQ386453.1  GI:4357476
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-154D12.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7558835"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-154D12"
                     /sex="Male"
                     /cell_type="Lymphocytes"
```

```
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

  Query Match             53.7%;  Score 27.4;  DB 28;  Length 700;
  Best Local Similarity   75.6%;  Pred. No. 2.3e+02;
  Matches   34;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAA 45
              | | | | |||||||||||| ||   || ||||| |||||| ||||
Db        545 ATATAGATAAATTGGACTTCAATAAAATTACTTGTGTGCTGCAAA 501


RESULT 15
AQ887236
LOCUS       AQ887236                449 bp    DNA     linear   GSS 10-NOV-1999
DEFINITION  HS_5552_B2_E03_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=9320 Col=6 Row=J, genomic survey sequence.
ACCESSION   AQ887236
VERSION     AQ887236.1  GI:6343522
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 9320  row: J  column: 6
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=9320 Col=6 Row=J"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

  Query Match             53.3%;  Score 27.2;  DB 28;  Length 449;
  Best Local Similarity   78.0%;  Pred. No. 2.8e+02;
  Matches   32;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              | || | | || |||||| |||||| | | ||||||||||||
Db        253 ATAACAGACTGAACTTAATGTTAAAAAATNTTGTGCTTCAA 293


Search completed: May  7, 2004, 15:42:23
Job time : 943.083 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:31:30 ; Search time 105.283 Seconds
(without alignments)
2194.362 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883171

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		38.4	75.3	13249	15	US-10-311-455-89	Sequence 89, Appl
2	c	38.4	75.3	13249	15	US-10-311-455-90	Sequence 90, Appl
3		29.8	58.4	605	13	US-10-027-632-224858	Sequence 224858,
4		29.8	58.4	605	16	US-10-027-632-224858	Sequence 224858,
5	c	28.6	56.1	879	13	US-10-282-122A-27417	Sequence 27417, A
6	c	28.6	56.1	10809	10	US-09-960-870-7	Sequence 7, Appli
7	c	28.6	56.1	10809	10	US-09-960-858-7	Sequence 7, Appli
8	c	28.6	56.1	10809	13	US-10-251-668-7	Sequence 7, Appli
9		28.6	56.1	580073	15	US-10-205-220-1	Sequence 1, Appli
10		26.2	51.4	356	13	US-10-424-599-8350	Sequence 8350, Ap
11	c	26.2	51.4	493	13	US-10-027-632-323533	Sequence 323533,
12	c	26.2	51.4	493	16	US-10-027-632-323533	Sequence 323533,
13	c	26.2	51.4	1371	13	US-10-027-632-103827	Sequence 103827,
14	c	26.2	51.4	1371	16	US-10-027-632-103827	Sequence 103827,
15		26	51.0	45698	11	US-09-984-429-344	Sequence 344, App
16	c	25.8	50.6	105219	13	US-10-087-192-658	Sequence 658, App
17		25.6	50.2	598	13	US-10-027-632-235079	Sequence 235079,
18		25.6	50.2	598	16	US-10-027-632-235079	Sequence 235079,
19	c	25.6	50.2	5864	15	US-10-311-455-1414	Sequence 1414, Ap
20	c	25.6	50.2	5864	15	US-10-240-452-62	Sequence 62, Appl
21	c	25.6	50.2	822900	16	US-10-292-798-1393	Sequence 1393, Ap
22		25.2	49.4	347	13	US-10-085-783A-31332	Sequence 31332, A
23		25.2	49.4	347	16	US-10-242-535A-31332	Sequence 31332, A
24	c	25.2	49.4	1214	14	US-10-042-417-45	Sequence 45, Appl
25	c	25.2	49.4	1285	9	US-09-764-847-268	Sequence 268, App
26	c	25.2	49.4	1285	15	US-10-092-154-268	Sequence 268, App
27	c	25.2	49.4	18434	15	US-10-311-455-1979	Sequence 1979, Ap
28		25	49.0	333	13	US-10-027-632-36924	Sequence 36924, A
29		25	49.0	333	16	US-10-027-632-36924	Sequence 36924, A
30		25	49.0	446	13	US-10-027-632-309812	Sequence 309812,
31		25	49.0	446	16	US-10-027-632-309812	Sequence 309812,
32	c	25	49.0	9888	15	US-10-311-455-1213	Sequence 1213, Ap
33		25	49.0	378361	10	US-09-901-136-3	Sequence 3, Appli
34	c	24.6	48.2	561	13	US-10-027-632-322541	Sequence 322541,
35	c	24.6	48.2	561	16	US-10-027-632-322541	Sequence 322541,
36		24.6	48.2	726	15	US-10-081-051-75	Sequence 75, Appl
37		24.6	48.2	948	9	US-09-974-300-57	Sequence 57, Appl
38		24.6	48.2	1018	10	US-09-933-767-15	Sequence 15, Appl
39		24.6	48.2	1018	13	US-10-004-860-15	Sequence 15, Appl
40		24.6	48.2	1018	15	US-10-023-282-15	Sequence 15, Appl
41		24.6	48.2	1218	9	US-09-974-300-297	Sequence 297, App
42		24.6	48.2	4460	15	US-10-081-051-74	Sequence 74, Appl
43	c	24.6	48.2	5678	15	US-10-311-455-1111	Sequence 1111, Ap
44		24.6	48.2	22927	10	US-09-764-891-7470	Sequence 7470, Ap
45		24.6	48.2	87878	12	US-10-052-482-82	Sequence 82, Appl

ALIGNMENTS

RESULT 1
US-10-311-455-89
; Sequence 89, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 89
; LENGTH: 13249
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-89

Query Match 75.3%; Score 38.4; DB 15; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.059;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709
```

```
RESULT 2
US-10-311-455-90/c
; Sequence 90, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 90
;   LENGTH: 13249
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-90

  Query Match             75.3%;  Score 38.4;  DB 15;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.059;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  |||||||  |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 3
US-10-027-632-224858
; Sequence 224858, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 224858
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(605)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-224858

  Query Match             58.4%;  Score 29.8;  DB 13;  Length 605;
  Best Local Similarity   79.1%;  Pred. No. 12;
  Matches   34;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          9 AAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||  ||||||| ||| ||  |||||||||||| | | |||
Db        499 AAATTTAACTTAAAATTATATTNTTTTGTGCTTCANAGACCAT 541


RESULT 4
US-10-027-632-224858
; Sequence 224858, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 224858
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(605)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-224858

  Query Match             58.4%;  Score 29.8;  DB 16;  Length 605;
  Best Local Similarity   79.1%;  Pred. No. 12;
  Matches   34;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          9 AAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||  ||||||| ||| ||  |||||||||||| | | |||
Db        499 AAATTTAACTTAAAATTATATTNTTTTGTGCTTCANAGACCAT 541


RESULT 5
US-10-282-122A-27417/c
; Sequence 27417, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
```

```
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27417
;   LENGTH: 879
;   TYPE: DNA
;   ORGANISM: Mycoplasma genitalium
US-10-282-122A-27417

  Query Match             56.1%;  Score 28.6;  DB 13;  Length 879;
  Best Local Similarity   72.5%;  Pred. No. 32;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  ||||||||| | ||| ||   | ||
Db        537 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 487


RESULT 6
US-09-960-870-7/c
; Sequence 7, Application US/09960870
; Publication No. US20030134281A1
; GENERAL INFORMATION:
;  APPLICANT: Evans, Glen
;  TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: P-EA 4738
;  CURRENT APPLICATION NUMBER: US/09/960,870
;  CURRENT FILING DATE:  2001-09-20
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 10809
;   TYPE: DNA
;   ORGANISM: M. genitalium
US-09-960-870-7

  Query Match             56.1%;  Score 28.6;  DB 10;  Length 10809;
  Best Local Similarity   72.5%;  Pred. No. 59;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  ||||||||| | ||| ||   | ||
Db       4241 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 4191


RESULT 7
US-09-960-858-7/c
; Sequence 7, Application US/09960858
; Publication No. US20030138777A1
; GENERAL INFORMATION:
;  APPLICANT: Evans, Glen
;  TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: P-EA 4974
;  CURRENT APPLICATION NUMBER: US/09/960,858
;  CURRENT FILING DATE:  2001-09-20
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 10809
;   TYPE: DNA
;   ORGANISM: M. genitalium
US-09-960-858-7

  Query Match             56.1%;  Score 28.6;  DB 10;  Length 10809;
  Best Local Similarity   72.5%;  Pred. No. 59;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  ||||||||| | ||| ||   | ||
Db       4241 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 4191


RESULT 8
US-10-251-668-7/c
; Sequence 7, Application US/10251668
; Publication No. US20040063097A1
; GENERAL INFORMATION:
;  APPLICANT: Evans, Glen
;  TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: P-EA 5441
;  CURRENT APPLICATION NUMBER: US/10/251,668
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US 09/960,607
;  PRIOR FILING DATE: 2001-09-20
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 10809
;   TYPE: DNA
;   ORGANISM: M. genitalium
US-10-251-668-7

  Query Match             56.1%;  Score 28.6;  DB 13;  Length 10809;
  Best Local Similarity   72.5%;  Pred. No. 59;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  ||||||||| | ||| ||   | ||
Db       4241 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 4191


RESULT 9
US-10-205-220-1
; Sequence 1, Application US/10205220
; Publication No. US20030170663A1
; GENERAL INFORMATION:
;  APPLICANT: Fraser et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmer
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB193P1D1
;  CURRENT APPLICATION NUMBER: US/10/205,220
;  CURRENT FILING DATE:  2002-07-26
;  PRIOR APPLICATION NUMBER: US 08/545,528
;  PRIOR FILING DATE: 1995-10-19
;  PRIOR APPLICATION NUMBER: US 08/488,018
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/473,545
;  PRIOR FILING DATE: 1995-06-07
;  NUMBER OF SEQ ID NOS: 1
```

```
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 580073
;   TYPE: DNA
;   ORGANISM: Mycoplasma genitalium
US-10-205-220-1

  Query Match             56.1%;  Score 28.6;  DB 15;  Length 580073;
  Best Local Similarity   72.5%;  Pred. No. 1.6e+02;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | ||||| |||||||||| | ||| ||   | ||
Db     508476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 508526


RESULT 10
US-10-424-599-8350
; Sequence 8350, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 8350
;   LENGTH: 356
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_10754C.1
US-10-424-599-8350

  Query Match             51.4%;  Score 26.2;  DB 13;  Length 356;
  Best Local Similarity   72.3%;  Pred. No. 1.4e+02;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATC 49
              ||||| |||||  ||| ||| | |||  | |||   ||||||| |||
Db         33 AAAAAAAAATTTAACTAAAAATAAAAACCCTTTTCTCTTCAAATATC 79


RESULT 11
US-10-027-632-323533/c
; Sequence 323533, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 323533
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-323533

  Query Match             51.4%;  Score 26.2;  DB 13;  Length 493;
  Best Local Similarity   72.3%;  Pred. No. 1.5e+02;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              ||||| ||||| ||||||| | |  || | || |||  || ||| ||
Db        139 CAAAATGAAATGGGACTTATAATACAAAAGTTGTGTTTTTAAAAAAT 93


RESULT 12
US-10-027-632-323533/c
; Sequence 323533, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 323533
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-323533

  Query Match             51.4%;  Score 26.2;  DB 16;  Length 493;
  Best Local Similarity   72.3%;  Pred. No. 1.5e+02;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              ||||| ||||| ||||||| | |  || | || |||  || ||| ||
Db        139 CAAAATGAAATGGGACTTATAATACAAAAGTTGTGTTTTTAAAAAAT 93


RESULT 13
US-10-027-632-103827/c
; Sequence 103827, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
```

```
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103827
;   LENGTH: 1371
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103827

  Query Match             51.4%;  Score 26.2;  DB 13;  Length 1371;
  Best Local Similarity   72.3%;  Pred. No. 1.9e+02;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              ||||| ||||| ||||||| | |  || | || |||  || ||| ||
Db        139 CAAAATGAAATGGGACTTATAATACAAAAGTTGTGTTTTTAAAAAAT 93


RESULT 14
US-10-027-632-103827/c
; Sequence 103827, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103827
;   LENGTH: 1371
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103827

  Query Match             51.4%;  Score 26.2;  DB 16;  Length 1371;
  Best Local Similarity   72.3%;  Pred. No. 1.9e+02;
  Matches   34;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              ||||| ||||| ||||||| | |  || | || |||  || ||| ||
Db        139 CAAAATGAAATGGGACTTATAATACAAAAGTTGTGTTTTTAAAAAAT 93


RESULT 15
US-09-984-429-344
; Sequence 344, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 53 Human Secreted Proteins
;  FILE REFERENCE: PZ018P2
;  CURRENT APPLICATION NUMBER: US/09/984,429
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,591
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/288,143
;  PRIOR FILING DATE: 1999-04-08
;  PRIOR APPLICATION NUMBER: PCT/US98/21142
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/061,463
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,529
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/071,498
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,527
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,536
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,532
;  PRIOR FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 727
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
;   LENGTH: 45698
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-984-429-344

  Query Match             51.0%;  Score 26;  DB 11;  Length 45698;
  Best Local Similarity   70.0%;  Pred. No. 5.4e+02;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | | ||| |||||  || | ||| |||| ||||| |||||   ||
Db      11579 AAACAGATAAACTGGACTCCAAATCAAAAACTTCTGTGCATCAAAGGACA 11628


Search completed: May  7, 2004, 17:31:44
Job time : 109.283 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:35:03 ; Search time 25.6094 Seconds
(without alignments)
1105.159 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		28.6	56.1	580073	4	US-08-545-528D-1	Sequence 1, Appli
2		24.6	48.2	1018	4	US-09-205-258-15	Sequence 15, Appl
3	c	24.4	47.8	3921	2	US-08-567-375-3	Sequence 3, Appli
4	c	24.4	47.8	3921	2	US-08-587-680A-3	Sequence 3, Appli
5	c	24.4	47.8	5992	2	US-08-475-891A-3	Sequence 3, Appli
6		24	47.1	1856	4	US-09-205-258-52	Sequence 52, Appl
7		23.8	46.7	1011	4	US-09-976-594-1069	Sequence 1069, Ap
8	c	23.8	46.7	1312	4	US-09-976-594-886	Sequence 886, App
9		23.8	46.7	1664976	4	US-08-916-421B-1	Sequence 1, Appli
10	c	23.4	45.9	6609	4	US-09-976-594-690	Sequence 690, App
11	c	23.4	45.9	786431	4	US-09-751-389-3	Sequence 3, Appli
12		23.2	45.5	963	4	US-09-328-352-1552	Sequence 1552, Ap
13		23.2	45.5	1506	4	US-09-134-001C-1278	Sequence 1278, Ap
14	c	23	45.1	1026	4	US-09-540-236-1082	Sequence 1082, Ap
15	c	23	45.1	19988	4	US-09-596-002-10	Sequence 10, Appl
16	c	22.8	44.7	458	4	US-09-387-286-35	Sequence 35, Appl
17		22.8	44.7	5532	4	US-08-956-171E-530	Sequence 530, App
18	c	22.8	44.7	6256	2	US-08-475-891A-1	Sequence 1, Appli
19	c	22.8	44.7	6256	2	US-08-567-375-1	Sequence 1, Appli
20	c	22.8	44.7	6256	2	US-08-587-680A-1	Sequence 1, Appli
21		22.6	44.3	466	4	US-09-621-976-13701	Sequence 13701, A
22		22.6	44.3	561	4	US-09-601-198-172	Sequence 172, App
23		22.4	43.9	446	4	US-09-621-976-9662	Sequence 9662, Ap
24		22.4	43.9	966	4	US-09-107-532A-1038	Sequence 1038, Ap
25		22.4	43.9	1731	4	US-09-134-001C-1118	Sequence 1118, Ap
26		22.4	43.9	3885	4	US-09-328-352-2188	Sequence 2188, Ap
27		22.4	43.9	5895	4	US-08-956-171E-1	Sequence 1, Appli
28	c	22.4	43.9	36159	4	US-09-749-588-3	Sequence 3, Appli
29		22.2	43.5	813	4	US-09-134-001C-2724	Sequence 2724, Ap
30	c	22.2	43.5	1245	4	US-09-543-681A-3265	Sequence 3265, Ap
31		22.2	43.5	1281	4	US-09-134-000C-1305	Sequence 1305, Ap
32	c	22.2	43.5	1618	4	US-09-800-729-29	Sequence 29, Appl
33		22.2	43.5	1626	4	US-09-358-383C-6	Sequence 6, Appli
34		22.2	43.5	2694	4	US-09-358-383C-4	Sequence 4, Appli
35		22.2	43.5	3321	4	US-09-358-383C-17	Sequence 17, Appl
36		22.2	43.5	5107	4	US-09-358-383C-15	Sequence 15, Appl
37		22.2	43.5	5955	4	US-09-358-383C-14	Sequence 14, Appl
38		22.2	43.5	202001	4	US-09-734-674-3	Sequence 3, Appli
39		22	43.1	400	4	US-08-956-171E-2161	Sequence 2161, Ap
40	c	22	43.1	904	4	US-09-171-209-59	Sequence 59, Appl
41	c	22	43.1	1044	4	US-09-221-017B-113	Sequence 113, App
42		22	43.1	1137	4	US-09-134-001C-657	Sequence 657, App
43		22	43.1	1782	2	US-08-714-168-3	Sequence 3, Appli
44		22	43.1	1782	3	US-09-320-721A-3	Sequence 3, Appli
45	c	22	43.1	5238	4	US-08-961-527-150	Sequence 150, App

ALIGNMENTS

RESULT 1
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmer
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match 56.1%; Score 28.6; DB 4; Length 580073;
Best Local Similarity 72.5%; Pred. No. 1.4;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | ||||| |||||||||| | ||| ||   | ||
Db     508476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 508526
```

RESULT 2
US-09-205-258-15
; Sequence 15, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881

```
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 1018
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-205-258-15

  Query Match             48.2%;  Score 24.6;  DB 4;  Length 1018;
  Best Local Similarity   70.2%;  Pred. No. 15;
  Matches   33;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATC 49
              ||| ||| |||| | | ||| |||||| |||||| | || |||
Db        746 AAATAGATATTGCTCATTAAGGTAAATATTTTTGTTGATGAATGATC 792


RESULT 3
US-08-567-375-3/c
; Sequence 3, Application US/08567375
; Patent No. 5952485
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Szabo, Veronique
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/567,375
;      FILING DATE:  04-DEC-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,645
;      FILING DATE:  29-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/475,891
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  023070-058930
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3921 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(1..2676, 3520..3918)
;      OTHER INFORMATION:  /product= "Xa-21"
US-08-567-375-3
```

```
  Query Match             47.8%;  Score 24.4;  DB 2;  Length 3921;
  Best Local Similarity   68.0%;  Pred. No. 19;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | |||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3456 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3407


RESULT 4
US-08-587-680A-3/c
; Sequence 3, Application US/08587680A
; Patent No. 5977434
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Szabo, Veronique
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/587,680A
;      FILING DATE:  17-JAN-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/475,891
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,645
;      FILING DATE:  29-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/567,375
;      FILING DATE:  04-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02307O-058940US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3921 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(1..2676, 3520..3918)
;      OTHER INFORMATION:  /product= "Xa-21"
US-08-587-680A-3

  Query Match             47.8%;  Score 24.4;  DB 2;  Length 3921;
  Best Local Similarity   68.0%;  Pred. No. 19;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | |||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3456 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3407


RESULT 5
US-08-475-891A-3/c
; Sequence 3, Application US/08475891A
; Patent No. 5859339
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,891A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02370-058910US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5992 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(512..3149, 3993..4393)
;      OTHER INFORMATION:  /product= "RRK-B"
;      OTHER INFORMATION:  /note= "Xa21 Xanthomonas spp. disease
;      OTHER INFORMATION:  resistance gene RRK-B from rice (Oryza
;      OTHER INFORMATION:  sativa)"
US-08-475-891A-3

  Query Match             47.8%;  Score 24.4;  DB 2;  Length 5992;
  Best Local Similarity   68.0%;  Pred. No. 20;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | |||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3928 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3879


RESULT 6
US-09-205-258-52
; Sequence 52, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
```

```
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 1856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-205-258-52

  Query Match             47.1%;  Score 24;  DB 4;  Length 1856;
  Best Local Similarity   68.8%;  Pred. No. 24;
  Matches   33;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||| || ||  | |||||      |||| | | ||||||||||| ||
Db        205 AAAAGTAAATGTCCATAAAGCGTTTCACTTATATTCTTCAAACATGAT 252


RESULT 7
US-09-976-594-1069
; Sequence 1069, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 1069
;   LENGTH: 1011
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 107309.1
;   NAME/KEY: unsure
;   LOCATION: 433-436, 445, 447, 454, 456, 463-465, 472, 495, 498, 662, 939
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-1069

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1011;
  Best Local Similarity   72.1%;  Pred. No. 27;
  Matches   31;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              |||||||| |  ||| ||||| || |  | |  ||||||| ||
Db        258 CAAAAAGATAGAGGATTTAAATTTCACAATTGATGTGCTTTAA 300


RESULT 8
US-09-976-594-886/c
; Sequence 886, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
```

```
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 886
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 981037.1
US-09-976-594-886

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1312;
  Best Local Similarity   80.0%;  Pred. No. 27;
  Matches   28;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGT 37
              ||| ||||||| || || ||||||||| |||| |
Db       1218 AAAGAGAAATTTTACATATAGTTAAATAATTTTTT 1184


RESULT 9
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (674435)..(674435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (741684)..(741684)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1664976;
  Best Local Similarity   66.7%;  Pred. No. 50;
  Matches   34;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              | ||| | |||||   | |  |  || ||| |||  ||| |||||||| ||
Db     555193 AAAAAGAAAAATTAATCATGCAAATATATAATTTAATGCATCAAACATTAT 555243


RESULT 10
US-09-976-594-690/c
; Sequence 690, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 690
;   LENGTH: 6609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 175918.15
;   NAME/KEY: unsure
;   LOCATION: 825
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-690

  Query Match             45.9%;  Score 23.4;  DB 4;  Length 6609;
  Best Local Similarity   67.3%;  Pred. No. 43;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATC 49
              || |||| ||||||     |||| ||||    |||   |||||||||  |
Db       5196 ACCAAAATAAATTGTGAAAAAAGATAAAAGACTTTCATCTTCAAACAAC 5148


RESULT 11
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001067
;  CURRENT APPLICATION NUMBER: US/09/751,389
;  CURRENT FILING DATE:  2001-01-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 786431
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(786431)
;   OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

  Query Match             45.9%;  Score 23.4;  DB 4;  Length 786431;
  Best Local Similarity   81.8%;  Pred. No. 65;
  Matches   27;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         18 TTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||| ||||| |||||||||| ||||| | ||
Db     658324 TGAAACTTAAAAACTTTTGTGCCTCAAAAAACA 658292


RESULT 12
US-09-328-352-1552
; Sequence 1552, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
```

```
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1552
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-1552

  Query Match             45.5%;  Score 23.2;  DB 4;  Length 963;
  Best Local Similarity   70.5%;  Pred. No. 42;
  Matches   31;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              | |  ||  || ||||||||||  ||  | |||||| ||| ||
Db        263 ATATTAATGAAGTGGACTTAAATGTACGTCCTTTTGTTCTTAAA 306


RESULT 13
US-09-134-001C-1278
; Sequence 1278, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1278
;   LENGTH: 1506
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1278

  Query Match             45.5%;  Score 23.2;  DB 4;  Length 1506;
  Best Local Similarity   77.8%;  Pred. No. 43;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGT 37
              |||||||||| |||  ||||||| |||| | || | |
Db       1360 CAAAAAGAAACTGGTATTAAAGGTAAACAATTATTT 1395


RESULT 14
US-09-540-236-1082/c
; Sequence 1082, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
;  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.2005-001
;  CURRENT APPLICATION NUMBER: US/09/540,236
;  CURRENT FILING DATE: 2000-04-04
;  NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1082
;   LENGTH: 1026
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-1082

  Query Match             45.1%;  Score 23;  DB 4;  Length 1026;
  Best Local Similarity   68.1%;  Pred. No. 49;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||   |||  ||  ||||| |||| | ||||  ||| | ||| ||||
Db        394 AATCCAAAAAGGCTTTAAATTTAATTTCTTTGTTGCATAAAATATCA 348


RESULT 15
US-09-596-002-10/c
; Sequence 10, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
;  APPLICANT:      Lagace, Robert, E.
;  APPLICANT:  Patterson, Chandra
;  APPLICANT:  Berg, Kim, L.
;  TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
;  FILE REFERENCE:  PM-0008-4 US
;  CURRENT APPLICATION NUMBER: US/09/596,002
;  CURRENT FILING DATE:  2000-06-16
;  PRIOR APPLICATION NUMBER:  60/140,121
;  PRIOR FILING DATE:  1999-06-18
;  NUMBER OF SEQ ID NOS:      41
;  SOFTWARE:       PERL Program
; SEQ ID NO 10
;   LENGTH: 19988
;   TYPE: DNA
;   ORGANISM: M. catarrhalis
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte template ID No..6632636 10
;  PUBLICATION INFORMATION:
US-09-596-002-10

  Query Match             45.1%;  Score 23;  DB 4;  Length 19988;
  Best Local Similarity   68.1%;  Pred. No. 65;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||   |||  ||  ||||| |||| | ||||  ||| | ||| ||||
Db       1287 AATCCAAAAAGGCTTTAAATTTAATTTCTTTGTTGCATAAAATATCA 1241


Search completed: May  7, 2004, 15:44:33
Job time : 34.6094 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 11:56:28 ; Search time 104.189 Seconds
(without alignments)
2079.475 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747720

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	48	94.1	168174	6	ABT11173	Abt11173 Human 5-l
	2	48	94.1	168273	6	ABT11114	Abt11114 Human 5-l
	3	38.4	75.3	13249	6	ABL32116	Abl32116 Human imm
c	4	38.4	75.3	13249	6	ABL32117	Abl32117 Human imm
c	5	38.4	75.3	13249	6	ABK31177	Abk31177 Signal tr
	6	38.4	75.3	13249	6	ABK31176	Abk31176 Signal tr
c	7	38.4	75.3	13249	6	ABL70132	Abl70132 Chemicall
	8	38.4	75.3	13249	6	ABL70131	Abl70131 Chemicall
c	9	28.6	56.1	879	7	ACA39547	Aca39547 Prokaryot
c	10	28.6	56.1	10809	7	ACC69139	Acc69139 M. genita
	11	28.6	56.1	80073	2	AAT58840_5	Continuation (6 of
	12	28.6	56.1	110000	2	AAT58840_4	Continuation (5 of
c	13	26.4	51.8	355	4	AAI83146	Aai83146 Human pol
c	14	25.6	50.2	5864	6	ABL33441	Abl33441 Human imm
c	15	25.6	50.2	5864	6	ABL54362	Abl54362 Chemicall
c	16	25.6	50.2	349938	9	ADC87621	Adc87621 Human GPC
	17	25.4	49.8	607	5	ABV58521	Abv58521 Human pro
	18	25.2	49.4	206	4	AAK81112	Aak81112 Human imm
c	19	25.2	49.4	1214	3	AAZ93368	Aaz93368 Sequence
c	20	25.2	49.4	1214	6	AAL41059	Aal41059 cDNA of H
c	21	25.2	49.4	1285	4	ABK41870	Abk41870 cDNA enco
c	22	25.2	49.4	1285	8	ADB59537	Adb59537 Connectiv
	23	25.2	49.4	1779	3	AAC80574	Aac80574 Human sec
	24	25.2	49.4	5728	4	AAK81109	Aak81109 Human imm
	25	25.2	49.4	5733	4	AAK81111	Aak81111 Human imm
	26	25.2	49.4	5733	4	AAK81110	Aak81110 Human imm
c	27	25.2	49.4	18434	6	ABL34006	Abl34006 Human imm
c	28	25	49.0	4129	4	ABL17308	Abl17308 Drosophil
c	29	25	49.0	9888	6	ABL33240	Abl33240 Human imm
	30	25	49.0	110000	7	AAL52246_1	Continuation (2 of
c	31	25	49.0	240000	7	ACD13446	Acd13446 Human DNA
	32	24.6	48.2	446	4	AAI81019	Aai81019 Human pol
	33	24.6	48.2	948	6	ABK72766	Abk72766 Bacillus
	34	24.6	48.2	1018	2	AAV84415	Aav84415 Human sec
	35	24.6	48.2	1018	4	ABA83198	Aba83198 Human sec
	36	24.6	48.2	1018	8	ACH04699	Ach04699 Novel hum
	37	24.6	48.2	1018	8	ACD44509	Acd44509 Human cDN
	38	24.6	48.2	1218	6	ABK73006	Abk73006 Bacillus
	39	24.6	48.2	3982	4	AAK78949	Aak78949 Human imm
	40	24.6	48.2	4460	6	AAD48267	Aad48267 Ehrlichia
c	41	24.6	48.2	5678	6	ABL33138	Abl33138 Human imm
	42	24.6	48.2	22927	4	AAL04782	Aal04782 Human rep
	43	24.6	48.2	22927	4	ABL97677	Abl97677 Human tes
	44	24.6	48.2	87878	8	ADA02576	Ada02576 Human FKB
	45	24.6	48.2	87878	9	ADB72314	Adb72314 Human FKB

ALIGNMENTS

RESULT 1
ABT11173
ID ABT11173 standard; DNA; 168174 BP.
XX
AC ABT11173;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 63.
XX
KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200262825-A2.
XX
PD 15-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-US003546.
XX
PR 08-FEB-2001; 2001US-0267515P.
PR 21-AUG-2001; 2001US-0314248P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes G, Meyer J;
XX
DR WPI; 2002-627522/67.
XX
PT New isolated nucleic acid molecule with an allelic variant of a
PT polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT prognosticating disorders associated with an aberrant inflammatory
PT response such as asthma.
XX
PS Disclosure; Fig 4; 290pp; English.
XX
CC The invention relates to an isolated human nucleic acid molecule
CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC selected from any of 3, 20 or 21 base pair sequences, given in the
CC specification, or their complement. The compositions and methods of the

```
CC   present invention are useful for diagnosing and/or prognosing disorders
CC   associated with an aberrant inflammatory response such as asthma,
CC   bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC   rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC   rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC   disease, atopic and contact dermatitis. The nucleic acid molecules can
CC   also be useful for identifying an individual amongst other individuals
CC   from the same species for use in forensic medicine and paternity testing.
CC   This polynucleotide sequence represents DNA relating to the human 5-
CC   lipoxygenase (5-LO) gene of the invention
XX
SQ   Sequence 168174 BP; 46808 A; 36442 C; 36942 G; 46474 T; 0 U; 1508 Other;

  Query Match             94.1%;  Score 48;  DB 6;  Length 168174;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-06;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166822 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 166869


RESULT 2
ABT11114
ID   ABT11114 standard; DNA; 168273 BP.
XX
AC   ABT11114;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Human 5-lipoxygenase gene related DNA sequence SEQ ID No 2.
XX
KW   Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW   sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW   rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW   polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW   atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200262825-A2.
XX
PD   15-AUG-2002.
XX
PF   07-FEB-2002; 2002WO-US003546.
XX
PR   08-FEB-2001; 2001US-0267515P.
PR   21-AUG-2001; 2001US-0314248P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Barnes G,  Meyer J;
XX
DR   WPI; 2002-627522/67.
XX
PT   New isolated nucleic acid molecule with an allelic variant of a
PT   polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT   prognosticating disorders associated with an aberrant inflammatory
PT   response such as asthma.
XX
PS   Disclosure; Fig 2; 290pp; English.
XX
CC   The invention relates to an isolated human nucleic acid molecule
CC   comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC   (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC   selected from any of 3, 20 or 21 base pair sequences, given in the
CC   specification, or their complement. The compositions and methods of the
CC   present invention are useful for diagnosing and/or prognosing disorders
CC   associated with an aberrant inflammatory response such as asthma,
CC   bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC   rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC   rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC   disease, atopic and contact dermatitis. The nucleic acid molecules can
CC   also be useful for identifying an individual amongst other individuals
CC   from the same species for use in forensic medicine and paternity testing.
CC   This polynucleotide sequence represents DNA relating to the human 5-
CC   lipoxygenase (5-LO) gene of the invention
XX
SQ   Sequence 168273 BP; 46834 A; 36467 C; 36966 G; 46498 T; 0 U; 1508 Other;

  Query Match             94.1%;  Score 48;  DB 6;  Length 168273;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-06;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     166921 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 166968


RESULT 3
ABL32116
ID   ABL32116 standard; DNA; 13249 BP.
XX
AC   ABL32116;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 89.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 89; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||| |||||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709
```

RESULT 4
ABL32117/c
ID ABL32117 standard; DNA; 13249 BP.
XX
AC ABL32117;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 90.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 90; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

```
  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||  |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541
```

RESULT 5
ABK31177/c
ID ABK31177 standard; DNA; 13249 BP.
XX
AC ABK31177;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #10.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007472.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction.
XX
PS Claim 1; SEQ ID NO 20; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC the cytosine methylation state (CpG islands) of these genes, and a method
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC genes associated with signal transduction. The genomic DNA can be
CC obtained from cells or cellular components which contain DNA, e.g. cell
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC brain, heart, prostate, lung, breast or liver, histologic object slides,
CC and all their possible combinations. The sequences of the invention are
CC useful for the diagnosis and therapy of diseases associated with signal
CC transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC chemically pretreated genomic DNA sequences of different genes associated
CC with signal transduction, or their complementary sequences. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

```
  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||  |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541
```

RESULT 6
ABK31176
ID ABK31176 standard; DNA; 13249 BP.
XX
AC ABK31176;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified DNA #10.

XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007472.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction.
XX
PS Claim 1; SEQ ID NO 19; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC the cytosine methylation state (CpG islands) of these genes, and a method
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC genes associated with signal transduction. The genomic DNA can be
CC obtained from cells or cellular components which contain DNA, e.g. cell
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC brain, heart, prostate, lung, breast or liver, histologic object slides,
CC and all their possible combinations. The sequences of the invention are
CC useful for the diagnosis and therapy of diseases associated with signal
CC transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC chemically pretreated genomic DNA sequences of different genes associated
CC with signal transduction, or their complementary sequences. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

Query Match 75.3%; Score 38.4; DB 6; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.0082;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
||||||||||||| ||||||||||| ||||||| || ||| || ||
Db 3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709

RESULT 7
ABL70132/c
ID ABL70132 standard; DNA; 13249 BP.
XX
AC ABL70132;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence complementary to#11.
XX
KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007471.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 22; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

Query Match 75.3%; Score 38.4; DB 6; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.0082;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
|||| ||||| ||||||| |||||||||||| | ||||||||||||
Db 9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541

RESULT 8
ABL70131
ID ABL70131 standard; DNA; 13249 BP.
XX
AC ABL70131;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence#11.
XX
KW Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007471.
XX
PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 21; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

Query Match 75.3%; Score 38.4; DB 6; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.0082;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709
```

RESULT 9
ACA39547/c
ID ACA39547 standard; DNA; 879 BP.
XX
AC ACA39547;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #21204.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycoplasma genitalium.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU35677.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 27417; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 879 BP; 269 A; 120 C; 143 G; 347 T; 0 U; 0 Other;

Query Match 56.1%; Score 28.6; DB 7; Length 879;
Best Local Similarity 72.5%; Pred. No. 8.8;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | ||||| |||||||||| | ||| ||   | ||
Db        537 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 487
```

RESULT 10
ACC69139/c
ID ACC69139 standard; DNA; 10809 BP.
XX
AC ACC69139;
XX
DT 10-JUL-2003 (first entry)
XX
DE M. genitalium aerobic metabolism gene cassette DNA SEQ ID NO:7.
XX
KW Mycoplasma genitalium; gene cassette; replication; transcription;
KW translation; metabolism; basic genetic operating system; gene therapy;
KW autonomous prototrophic nanomachine; auxotrophic nanomachine;
KW nanomachine; bioreactor; bioremediation; therapeutic; delivery system;
KW artificial tissue; artificial organ system; energy conversion system;
KW processing system; anabolic system; catabolic system; biological film;
KW biological coating; cosmetic; gene; ds.
XX
OS Mycoplasma genitalium.
XX

```
PN   WO2003025145-A2.
XX
PD   27-MAR-2003.
XX
PF   18-SEP-2002; 2002WO-US029811.
XX
PR   20-SEP-2001; 2001US-00960870.
XX
PA   (EGEA-) EGEA BIOSCIENCES INC.
XX
PI   Evans GA;
XX
DR   WPI; 2003-354602/33.
XX
PT   New basic genetic operating system for autonomous prototrophic or
PT   auxotrophic nanomachine, useful for therapeutic, diagnostic or industrial
PT   purposes, comprises a nanomachine genome encoding a gene set for
PT   viability or replication.
XX
PS   Example 1; Page 210-213; 250pp; English.
XX
CC   The present invention describes a basic genetic operating system for an
CC   autonomous prototrophic or auxotrophic nanomachine comprising a
CC   nanomachine genome encoding a minimal gene set sufficient for viability
CC   or replication, optionally in the presence of an auxotrophic molecule.
CC   Also described is an autonomous prototrophic or auxotrophic nanomachine
CC   comprising a basic genetic operating system for autonomous prototrophic
CC   or auxotrophic viability or replication, optionally in the presence of an
CC   auxotrophic molecule, and a particle envelope. The nanomachines can be
CC   used in gene therapy. The basic genetic operating system or nanomachine
CC   is useful in therapeutic, diagnostic and industrial applications, e.g. as
CC   a bioreactor, for bioremediation, for the production of a therapeutic
CC   biomolecule or as a therapeutic reagent, for the production of a
CC   diagnostic indicator or reagent, as a delivery system, as an artificial
CC   tissue or organ system, as an energy conversion system, as a processing
CC   system, as an anabolic or catabolic system, for the production of
CC   biological films or coatings, and for cosmetic applications. The present
CC   sequence represents a Mycoplasma genitalium gene cassette nucleotide
CC   sequence, which is used in an example from the present invention for the
CC   design and synthesis of a basic genetic operation system for a
CC   replication competent nanomachine
XX
SQ   Sequence 10809 BP; 3805 A; 1468 C; 1932 G; 3604 T; 0 U; 0 Other;

  Query Match             56.1%;  Score 28.6;  DB 7;  Length 10809;
  Best Local Similarity   72.5%;  Pred. No. 10;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  |||||||||| | ||| ||   | ||
Db       4241 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 4191


RESULT 11
AAT58840_5
Continuation (6 of 6) of AAT58840 from base 500001 (Mycoplasma genitalium genome. )
WP Sequence split into 6 fragments  LOCUS AAT58840 Accession Aat58840
WP      Fragment Name       Begin      End
WP      AAT58840_0              1      110000
WP      AAT58840_1         100001      210000
WP      AAT58840_2         200001      310000
WP      AAT58840_3         300001      410000
WP      AAT58840_4         400001      510000
WP      AAT58840_5         500001      580073

  Query Match             56.1%;  Score 28.6;  DB 2;  Length 80073;
  Best Local Similarity   72.5%;  Pred. No. 12;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  |||||||||| | ||| ||   | ||
Db       8476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 8526


RESULT 12
AAT58840_4
Continuation (5 of 6) of AAT58840 from base 400001 (Mycoplasma genitalium genome. )
WP Sequence split into 6 fragments  LOCUS AAT58840 Accession Aat58840
WP      Fragment Name       Begin      End
WP      AAT58840_0              1      110000
WP      AAT58840_1         100001      210000
WP      AAT58840_2         200001      310000
WP      AAT58840_3         300001      410000
WP      AAT58840_4         400001      510000
WP      AAT58840_5         500001      580073

  Query Match             56.1%;  Score 28.6;  DB 2;  Length 110000;
  Best Local Similarity   72.5%;  Pred. No. 12;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | |||||  |||||||||| | ||| ||   | ||
Db     108476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 108526


RESULT 13
AAI83146/c
ID   AAI83146 standard; cDNA; 355 BP.
XX
AC   AAI83146;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3206.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO03215.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 1; SEQ ID NO 3206; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
```

```
XX
SQ   Sequence 355 BP; 86 A; 65 C; 74 G; 130 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 26.4;  DB 4;  Length 355;
  Best Local Similarity   75.0%;  Pred. No. 41;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 AGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||||||||||||||| || || ||| |||  ||| |||   ||
Db        209 ATAAATTGGACTTAAAATTCAAAACTCTTGCCCTTTAAAAGACA 166


RESULT 14
ABL33441/c
ID   ABL33441 standard; DNA; 5864 BP.
XX
AC   ABL33441;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1414.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 1414; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 5864 BP; 1601 A; 136 C; 1351 G; 2776 T; 0 U; 0 Other;

  Query Match             50.2%;  Score 25.6;  DB 6;  Length 5864;
  Best Local Similarity   70.8%;  Pred. No. 89;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              | ||||| ||||   || ||||   ||| |||| | | ||||||| ||
Db       4241 AAAAAAACAAATAAAACCTAAAAACAAAAACTTATTTTCTTCAAAAAT 4194


RESULT 15
ABL54362/c
ID   ABL54362 standard; DNA; 5864 BP.
XX
AC   ABL54362;
XX
DT   29-JUL-2002  (first entry)
XX
DE   Chemically treated apoptosis gene complementary to gene #31.
XX
KW   Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;
KW   Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;
KW   cancer; ds.
XX
OS   Unidentified.
XX
PN   WO200177164-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-EP003969.
XX
PR   06-APR-2000; 2000DE-01019058.
PR   07-APR-2000; 2000DE-01019173.
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-017444/02.
XX
PT   Chemically modified sequences of genes associated with apoptosis are
PT   useful to determine methylation patterns of genomic DNA samples for
PT   diagnosis of associated diseases such as cancer.
XX
PS   Claim 1; Seq ID #62; 24pp; English.
XX
CC   This invention relates to chemically pre-treated DNA of genes associated
CC   with apoptosis. The nucleic acids are used to allocate patients for
CC   specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,
CC   neurodegenerative disorders, Herpes simplex virus infection, renal
CC   ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This
CC   nucleotide sequence represents a chemically treated apoptosis gene. Even
CC   SEQ ID numbers are the complementary DNA strands to the odd SEQ ID
CC   numbers. The sequence data for this patent is not represented in the
CC   printed specification but is based on information supplied by the
CC   European patent office
XX
SQ   Sequence 5864 BP; 1601 A; 136 C; 1351 G; 2776 T; 0 U; 0 Other;

  Query Match             50.2%;  Score 25.6;  DB 6;  Length 5864;
  Best Local Similarity   70.8%;  Pred. No. 89;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              | ||||| ||||   || ||||   ||| |||| | | ||||||| ||
Db       4241 AAAAAAACAAATAAAACCTAAAAACAAAAACTTATTTTCTTCAAAAAT 4194


Search completed: May  7, 2004, 13:50:20
Job time : 111.189 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:24:59 ; Search time 531.669 Seconds
(without alignments)
4157.648 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940541

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		48	94.1	129266	9	AL731567	AL731567 Human DNA
2		48	94.1	160654	2	AC011879	AC011879 Homo sapi
3		48	94.1	194453	2	AC010862	AC010862 Homo sapi
4		38.4	75.3	13249	6	AX344172	AX344172 Sequence
5	c	38.4	75.3	13249	6	AX344173	AX344173 Sequence
6		38.4	75.3	13249	6	AX345018	AX345018 Sequence
7	c	38.4	75.3	13249	6	AX345019	AX345019 Sequence
8		38.4	75.3	13249	6	AX348563	AX348563 Sequence
9	c	38.4	75.3	13249	6	AX348564	AX348564 Sequence
10	c	34	66.7	192044	9	AL590439	AL590439 Human DNA
11	c	33.6	65.9	128529	2	AC025758	AC025758 Homo sapi
12		33.6	65.9	157325	9	AC008810	AC008810 Homo sapi
13		33.6	65.9	164217	9	AC093264	AC093264 Homo sapi
14	c	33.4	65.5	209016	9	BS000239	BS000239 Pan trogl
15	c	32.6	63.9	173053	10	AL365334	AL365334 Mouse DNA
16	c	31.8	62.4	154157	9	HS101D08	AL133492 Homo sapi
17	c	31.8	62.4	223201	9	HS53I10	AL133493 Homo sapi
18	c	31.8	62.4	283388	2	AC012285	AC012285 Homo sapi
19	c	31.8	62.4	340000	9	HS21C103	AL163303 Homo sapi
20	c	30.6	60.0	60812	2	AC145991	AC145991 Pan trogl
21		30.6	60.0	82419	9	AC004979	AC004979 Homo sapi
22	c	30.6	60.0	133691	9	AC074347	AC074347 Homo sapi
23	c	30.2	59.2	58250	2	AC103690	AC103690 Homo sapi
24		30.2	59.2	144017	2	AF235106	AF235106 Homo sapi
25		30.2	59.2	169537	9	AC100814	AC100814 Homo sapi
26	c	30.2	59.2	175910	2	AC091004	AC091004 Homo sapi
27	c	30.2	59.2	181907	2	AC108731	AC108731 Homo sapi
28		29.8	58.4	160307	9	AC018359	AC018359 Homo sapi
29		29.8	58.4	171347	9	AC099776	AC099776 Homo sapi
30		29.4	57.6	174873	9	AC009069	AC009069 Homo sapi
31	c	29	56.9	83250	9	HS377F16	Z93783 Human DNA s
32		29	56.9	143878	2	AL359974	AL359974 Homo sapi
33	c	29	56.9	146250	2	AC074240	AC074240 Homo sapi
34		29	56.9	159231	9	AL161654	AL161654 Human DNA
35		29	56.9	160629	9	AC073326	AC073326 Homo sapi
36		29	56.9	168608	2	AL591477	AL591477 Homo sapi
37		29	56.9	179876	9	AC087863	AC087863 Homo sapi
38		29	56.9	191830	2	AC026332	AC026332 Homo sapi
39		28.8	56.5	202495	9	CNS01DW6	AL136418 Human chr
40		28.8	56.5	202496	9	CNS01DX6	AL139054 Human chr
41		28.6	56.1	9374	1	U39722	U39722 Mycoplasma
42		28.6	56.1	80073	6	AR300198_5	Continuation (6 of
43		28.6	56.1	87790	9	AC090698	AC090698 Homo sapi
44		28.6	56.1	110000	6	AR300198_4	Continuation (5 of
45	c	28.6	56.1	161198	2	AC015867	AC015867 Homo sapi

ALIGNMENTS

RESULT 1
AL731567
LOCUS AL731567 129266 bp DNA linear PRI 20-JUN-2002
DEFINITION Human DNA sequence from clone RP11-67C2 on chromosome 10, complete sequence.
ACCESSION AL731567 AC010865
VERSION AL731567.6 GI:21537524
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 129266)
AUTHORS Whitehead,S.
TITLE Direct Submission

```
  JOURNAL   Submitted (31-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 21, 2002 this sequence version replaced gi:21213582.
            Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver
            Street, Waltham, MA 02453, USA
            http://www.genomecorp.com
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-67C2 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .129266
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-67C2"
                     /clone_lib="RPCI-11.1"
ORIGIN

  Query Match             94.1%;  Score 48;  DB 9;  Length 129266;
  Best Local Similarity   100.0%;  Pred. No. 0.00039;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33190 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 33237


RESULT 2
AC011879
LOCUS       AC011879               160654 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-16P14, WORKING DRAFT SEQUENCE, 30 unordered
            pieces.
ACCESSION   AC011879
VERSION     AC011879.3  GI:7239554
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-16P14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 14, 2000 this sequence version replaced gi:6524208.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3606
                Center clone name: 16_P_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 111055 bases at least Q40
                Consensus quality: 135066 bases at least Q30
                Consensus quality: 147921 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 157754; sum-of-contigs
                Quality coverage: 2.9 in Q20 bases; agarose-fp
                Quality coverage: 3.0 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 30 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      151: contig of 151 bp in length
            *      152      251: gap of 100 bp
            *      252     1760: contig of 1509 bp in length
            *     1761     1860: gap of 100 bp
            *     1861     3069: contig of 1209 bp in length
            *     3070     3169: gap of 100 bp
            *     3170     4720: contig of 1551 bp in length
            *     4721     4820: gap of 100 bp
            *     4821     6174: contig of 1354 bp in length
            *     6175     6274: gap of 100 bp
            *     6275     7417: contig of 1143 bp in length
            *     7418     7517: gap of 100 bp
            *     7518     9158: contig of 1641 bp in length
            *     9159     9258: gap of 100 bp
            *     9259    10865: contig of 1607 bp in length
            *    10866    10965: gap of 100 bp
            *    10966    12859: contig of 1894 bp in length
            *    12860    12959: gap of 100 bp
            *    12960    15671: contig of 2712 bp in length
            *    15672    15771: gap of 100 bp
            *    15772    18082: contig of 2311 bp in length
            *    18083    18182: gap of 100 bp
            *    18183    20523: contig of 2341 bp in length
            *    20524    20623: gap of 100 bp
            *    20624    22903: contig of 2280 bp in length
            *    22904    23003: gap of 100 bp
            *    23004    23671: contig of 668 bp in length
            *    23672    23771: gap of 100 bp
```

```
            *    23772    25541: contig of 1770 bp in length
            *    25542    25641: gap of 100 bp
            *    25642    28323: contig of 2682 bp in length
            *    28324    28423: gap of 100 bp
            *    28424    31498: contig of 3075 bp in length
            *    31499    31598: gap of 100 bp
            *    31599    36626: contig of 5028 bp in length
            *    36627    36726: gap of 100 bp
            *    36727    42109: contig of 5383 bp in length
            *    42110    42209: gap of 100 bp
            *    42210    48339: contig of 6130 bp in length
            *    48340    48439: gap of 100 bp
            *    48440    55333: contig of 6894 bp in length
            *    55334    55433: gap of 100 bp
            *    55434    63594: contig of 8161 bp in length
            *    63595    63694: gap of 100 bp
            *    63695    73943: contig of 10249 bp in length
            *    73944    74043: gap of 100 bp
            *    74044    83665: contig of 9622 bp in length
            *    83666    83765: gap of 100 bp
            *    83766    95322: contig of 11557 bp in length
            *    95323    95422: gap of 100 bp
            *    95423   108403: contig of 12981 bp in length
            *   108404   108503: gap of 100 bp
            *   108504   120220: contig of 11717 bp in length
            *   120221   120320: gap of 100 bp
            *   120321   132958: contig of 12638 bp in length
            *   132959   133058: gap of 100 bp
            *   133059   145697: contig of 12639 bp in length
            *   145698   145797: gap of 100 bp
            *   145798   160654: contig of 14857 bp in length.
FEATURES             Location/Qualifiers
     source          1. .160654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-16P14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .151
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    252. .1760
                     /note="assembly_fragment"
     misc_feature    1861. .3069
                     /note="assembly_fragment"
     misc_feature    3170. .4720
                     /note="assembly_fragment"
     misc_feature    4821. .6174
                     /note="assembly_fragment"
     misc_feature    6275. .7417
                     /note="assembly_fragment"
     misc_feature    7518. .9158
                     /note="assembly_fragment"
     misc_feature    9259. .10865
                     /note="assembly_fragment"
     misc_feature    10966. .12859
                     /note="assembly_fragment"
     misc_feature    12960. .15671
                     /note="assembly_fragment"
     misc_feature    15772. .18082
                     /note="assembly_fragment"
     misc_feature    18183. .20523
                     /note="assembly_fragment"
     misc_feature    20624. .22903
                     /note="assembly_fragment"
     misc_feature    23004. .23671
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    23772. .25541
                     /note="assembly_fragment"
     misc_feature    25642. .28323
                     /note="assembly_fragment"
     misc_feature    28424. .31498
                     /note="assembly_fragment"
     misc_feature    31599. .36626
                     /note="assembly_fragment"
     misc_feature    36727. .42109
                     /note="assembly_fragment"
     misc_feature    42210. .48339
                     /note="assembly_fragment"
     misc_feature    48440. .55333
                     /note="assembly_fragment"
     misc_feature    55434. .63594
                     /note="assembly_fragment"
     misc_feature    63695. .73943
                     /note="assembly_fragment"
     misc_feature    74044. .83665
                     /note="assembly_fragment"
     misc_feature    83766. .95322
                     /note="assembly_fragment"
     misc_feature    95423. .108403
                     /note="assembly_fragment"
     misc_feature    108504. .120220
                     /note="assembly_fragment"
     misc_feature    120321. .132958
                     /note="assembly_fragment"
     misc_feature    133059. .145697
                     /note="assembly_fragment"
     misc_feature    145798. .160654
                     /note="assembly_fragment"
ORIGIN

  Query Match             94.1%;  Score 48;  DB 2;  Length 160654;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40267 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 40314


RESULT 3
AC010862
LOCUS       AC010862              194453 bp    DNA     linear   HTG 30-AUG-2001
DEFINITION  Homo sapiens chromosome 06 clone RP11-326D18, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC010862
VERSION     AC010862.7  GI:9957987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 194453)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 194453)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Sep 1, 2000 this sequence version replaced gi:8247773.
            -------------- Genome Center
                           Center: Genome Therapeutics Corperation
                           Center code: GTC
                           Web site: http://www.genomecorp.com/
                           Contact: gtc-seqcenter@genomecorp.com
                       -------------- Project Information
                           Center project name: hg024
                       -------------- Summary Statistics
                           Sequencing vector: N/A
                           Chemistry: Dye-terminator Big Dye; 100% of reads
```

```
                     Assembly program: Phrap; version 990315
                     Consensus quality: 162991 bases at least Q40
                     Consensus quality: 176452 bases at least Q30
                     Consensus quality: 179870 bases at least Q20
                     Insert size: 192053; sum-of-contigs
                     Quality coverage: 4.2x in Q20 bases; sum-of-contigs
                   --------------.
                * NOTE: This is a 'working draft' sequence. It currently
                * consists of 25 contigs. The true order of the pieces
                * is not known and their order in this sequence record is
                * arbitrary. Gaps between the contigs are represented as
                * runs of N, but the exact sizes of the gaps are unknown.
                * This record will be updated with the finished sequence
                * as soon as it is available and the accession number will
                * be preserved.
                *        1     1117: contig of 1117 bp in length
                *     1118     1217: gap of unknown length
                *     1218     2365: contig of 1148 bp in length
                *     2366     2465: gap of unknown length
                *     2466     3576: contig of 1111 bp in length
                *     3577     3676: gap of unknown length
                *     3677     4782: contig of 1106 bp in length
                *     4783     4882: gap of unknown length
                *     4883     6049: contig of 1167 bp in length
                *     6050     6149: gap of unknown length
                *     6150     7670: contig of 1521 bp in length
                *     7671     7770: gap of unknown length
                *     7771     9521: contig of 1751 bp in length
                *     9522     9621: gap of unknown length
                *     9622    10895: contig of 1274 bp in length
                *    10896    10995: gap of unknown length
                *    10996    12183: contig of 1188 bp in length
                *    12184    12283: gap of unknown length
                *    12284    13602: contig of 1319 bp in length
                *    13603    13702: gap of unknown length
                *    13703    15955: contig of 2253 bp in length
                *    15956    16055: gap of unknown length
                *    16056    18297: contig of 2242 bp in length
                *    18298    18397: gap of unknown length
                *    18398    24368: contig of 5971 bp in length
                *    24369    24468: gap of unknown length
                *    24469    30049: contig of 5581 bp in length
                *    30050    30149: gap of unknown length
                *    30150    36961: contig of 6812 bp in length
                *    36962    37061: gap of unknown length
                *    37062    43997: contig of 6936 bp in length
                *    43998    44097: gap of unknown length
                *    44098    49940: contig of 5843 bp in length
                *    49941    50040: gap of unknown length
                *    50041    56987: contig of 6947 bp in length
                *    56988    57087: gap of unknown length
                *    57088    65541: contig of 8454 bp in length
                *    65542    65641: gap of unknown length
                *    65642    75225: contig of 9584 bp in length
                *    75226    75325: gap of unknown length
                *    75326    85420: contig of 10095 bp in length
                *    85421    85520: gap of unknown length
                *    85521   101885: contig of 16365 bp in length
                *   101886   101985: gap of unknown length
                *   101986   124008: contig of 22023 bp in length
                *   124009   124108: gap of unknown length
                *   124109   157199: contig of 33091 bp in length
                *   157200   157299: gap of unknown length
                *   157300   194453: contig of 37154 bp in length.
FEATURES             Location/Qualifiers
     source          1. .194453
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="06"
                     /clone="RP11-326D18"
                     /clone_lib="RPCI-11"

ORIGIN
```

```
  Query Match             94.1%;  Score 48;  DB 2;  Length 194453;
  Best Local Similarity   100.0%;  Pred. No. 0.00035;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4527 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 4574


RESULT 4
AX344172
LOCUS       AX344172               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 19 from Patent WO0200926.
ACCESSION   AX344172
VERSION     AX344172.1  GI:18492060
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 19 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 5
AX344173/c
LOCUS       AX344173               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 20 from Patent WO0200926.
ACCESSION   AX344173
VERSION     AX344173.1  GI:18492061
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 20 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541
```

```
RESULT 6
AX345018
LOCUS       AX345018               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 89 from Patent WO0200928.
ACCESSION   AX345018
VERSION     AX345018.1  GI:18492904
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 89 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 7
AX345019/c
LOCUS       AX345019               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 90 from Patent WO0200928.
ACCESSION   AX345019
VERSION     AX345019.1  GI:18492905
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 90 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 8
AX348563
LOCUS       AX348563               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 21 from Patent WO0202807.
ACCESSION   AX348563
VERSION     AX348563.1  GI:18614598
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 21 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 9
AX348564/c
LOCUS       AX348564               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 22 from Patent WO0202807.
ACCESSION   AX348564
VERSION     AX348564.1  GI:18614599
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 22 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 10
AL590439/c
LOCUS       AL590439              192044 bp    DNA     linear   PRI 23-AUG-2001
DEFINITION  Human DNA sequence from clone RP11-394I23 on chromosome 10,
            complete sequence.
ACCESSION   AL590439
VERSION     AL590439.12  GI:15384822
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 192044)
  AUTHORS   Babbage,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
```

```
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 31, 2001 this sequence version replaced gi:14268248.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-394I23 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-394I23 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true right end of clone RP11-394I23 is at 192044 in this
            sequence. The true left end of clone RP11-657A9 is at 85254 in this
            sequence. The true right end of clone RP11-39J5 is at 100 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .192044
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-394I23"
                     /clone_lib="RPCI-11.2"
ORIGIN

  Query Match             66.7%;  Score 34;  DB 9;  Length 192044;
  Best Local Similarity   80.0%;  Pred. No. 2.6;
  Matches   40;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||||| | |||||||||||||| ||||| |||| |||| |||||   ||
Db      30825 ACAAAGATAAATTGGACTTAAAATTAAAAACTTGTGTGTGTCAAAGGACA 30776


RESULT 11
AC025758/c
LOCUS       AC025758               128529 bp    DNA     linear   HTG 18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone CTD-2235A13, WORKING DRAFT
            SEQUENCE, 16 ordered pieces.
ACCESSION   AC025758
VERSION     AC025758.3  GI:9256494
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 128529)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 128529)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:7711854.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 717802
            Center clone name: CITB-H1_2235A13
            -------
            Summary Statistics
            Consensus quality: 116698 bases at least Q40
            Consensus quality: 124565 bases at least Q30
            Consensus quality: 125980 bases at least Q20
            Estimated insert size: 130000; pulse field gel estimation
            Estimated insert size: 127829; sum-of-contigs estimation
            Quality coverage:  4.18 in Q20 bases; pulse field gel estimation
            Quality coverage:  4.25 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1     8719: contig of 8719 bp in length
            *    8720     8819: gap of unknown length
            *    8820    17381: contig of 8562 bp in length
            *   17382    17481: gap of unknown length
            *   17482    22126: contig of 4645 bp in length
            *   22127    22226: gap of unknown length
            *   22227    25476: contig of 3250 bp in length
            *   25477    25576: gap of unknown length
            *   25577    36711: contig of 11135 bp in length
            *   36712    36811: gap of unknown length
            *   36812    52089: contig of 15278 bp in length
            *   52090    52189: gap of unknown length
            *   52190    56295: contig of 4106 bp in length
            *   56296    56395: gap of unknown length
            *   56396    59531: contig of 3136 bp in length
            *   59532    59631: gap of unknown length
            *   59632    66671: contig of 7040 bp in length
            *   66672    66771: gap of unknown length
            *   66772    83163: contig of 16392 bp in length
            *   83164    83263: gap of unknown length
            *   83264   100033: contig of 16770 bp in length
            *  100034   100133: gap of unknown length
            *  100134   102212: contig of 2079 bp in length
            *  102213   102312: gap of unknown length
            *  102313   104754: contig of 2442 bp in length
            *  104755   104854: gap of unknown length
            *  104855   112328: contig of 7474 bp in length
            *  112329   112428: gap of unknown length
            *  112429   124204: contig of 11776 bp in length
            *  124205   124304: gap of unknown length
            *  124305   128529: contig of 4225 bp in length.
FEATURES             Location/Qualifiers
     source          1. .128529
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2235A13"
                     /clone_lib="CalTech human BAC library D"
ORIGIN

  Query Match             65.9%;  Score 33.6;  DB 2;  Length 128529;
  Best Local Similarity   81.2%;  Pred. No. 3.8;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||||  | |
Db       6736 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 6689


RESULT 12
AC008810
LOCUS       AC008810               157325 bp    DNA     linear   PRI 28-JUN-2003
DEFINITION  Homo sapiens chromosome 5 clone CTD-2096I23, complete sequence.
ACCESSION   AC008810
VERSION     AC008810.8  GI:32328942
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jun 28, 2003 this sequence version replaced gi:14993661.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.3.
FEATURES             Location/Qualifiers
     source          1. .157325
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2096I23"
ORIGIN

  Query Match           65.9%;  Score 33.6;  DB 9;  Length 157325;
  Best Local Similarity  81.2%;  Pred. No. 3.6;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||||  | |
Db     123706 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 123753


RESULT 13
AC093264
LOCUS       AC093264               164217 bp    DNA     linear   PRI 01-FEB-2002
DEFINITION  Homo sapiens chromosome 5 clone RP11-263G2, complete sequence.
ACCESSION   AC093264
VERSION     AC093264.2  GI:18464051
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:15193398.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .164217
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-263G2"
ORIGIN

  Query Match           65.9%;  Score 33.6;  DB 9;  Length 164217;
  Best Local Similarity  81.2%;  Pred. No. 3.5;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||||  | |
Db      30833 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 30880


RESULT 14
BS000239/c
LOCUS       BS000239               209016 bp    DNA     linear   PRI 07-OCT-2003
DEFINITION  Pan troglodytes chromosome 22 clone:PTB-153E07, map 22, complete
            sequences.
ACCESSION   BS000239
VERSION     BS000239.1  GI:37537506
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   The Chimpanzee Chromosome 22 Sequencing Consortium.
  TITLE     DNA sequence of chimpanzee chromosome 22 and its evolutionary
            implications
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209016)
  AUTHORS   Tsai,S., Liu,T., Wu,K., Liao,T. and Hsiao,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2003) Shih-Feng Tsai, National Health Research
            Institutes (NHRI), Division of Molecular and Genomic Medicine; 128,
            Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan
            (E-mail:petsai@nhri.org.tw, URL:http://www.nhri.org.tw/,
            Tel:886-2-28267319, Fax:886-2-28200552)
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
```

```
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
            Center: National Yang Ming University Genome Research Center Center
            code: YMGC
            Web site: http://genome.ym.edu.tw/
            Contact: sequence@ym.edu.tw
            -------------- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: HI
            -------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly
            program: Phrap; version 0.990319
            Consensus quality:   207,750 bases at least Q40
            Consensus quality:   207,996 bases at least Q30
            Consensus quality:   208,014 bases at least Q20
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            --------------
            Source information:
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS145
            --------------
            Sequence Quality Assessment:
            This entry has been annotated with sequence
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than 1 error in
            10,000 bp.
            ----------------------
            Neighboring clones: PTB-152N20(left) and RP43-055A16(right).
FEATURES             Location/Qualifiers
     source          1. .209016
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="22"
                     /clone="PTB-153E07"
                     /clone_lib="PTB1 chimpanzee BAC"
     misc_feature    16459. .16463
                     /note="low quality region"
     misc_feature    43909. .44908
                     /note="gap containing unresolved di-nucleotide repeats,
                     (TG)n"
     misc_feature    45769. .45771
                     /note="low quality region"
     misc_feature    45776
                     /note="low quality region"
     misc_feature    45778
                     /note="low quality region"
     misc_feature    46232. .46235
                     /note="low quality region"
     misc_feature    46277. .46280
                     /note="low quality region"
     misc_feature    128655
                     /note="low quality region"
     misc_feature    128691
                     /note="low quality region"
ORIGIN

  Query Match             65.5%;  Score 33.4;  DB 9;  Length 209016;
  Best Local Similarity   78.4%;  Pred. No. 3.8;
  Matches   40;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              | | ||| ||||||| |||||| ||||| |||||||||||||  ||   |||
Db     123272 AAATAAATAAATTGGGCTTAAAATTAAAAACTTTTGTGCTTGCAAGGACAT 123222


RESULT 15
AL365334/c
LOCUS       AL365334              173053 bp    DNA     linear   ROD 29-JUN-2002
DEFINITION  Mouse DNA sequence from clone RP23-392F1 on chromosome 1, complete
            sequence.
ACCESSION   AL365334
VERSION     AL365334.13  GI:20068419
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 173053)
  AUTHORS   Blakey,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 7, 2002 this sequence version replaced gi:14141369.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-392F1 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            --------------.
FEATURES             Location/Qualifiers
     source          1. .173053
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="RP23-392F1"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             63.9%;  Score 32.6;  DB 10;  Length 173053;
  Best Local Similarity   80.9%;  Pred. No. 6.6;
  Matches   38;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACA 47
              |||| || ||| | ||||| ||||||||||||| | | |||||||
Db     140706 ACAACAATAAAATCGACTTTAAGTTAAATACTTCTTTACTTCAAAAA 140660

Search completed: May  7, 2004, 14:30:58
Job time : 538.003 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:35:03 ; Search time 25.6094 Seconds
(without alignments)
1105.159 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	28.6	56.1	580073	4	US-08-545-528D-1	Sequence 1, Appli
	2	24.6	48.2	1018	4	US-09-205-258-15	Sequence 15, Appl
c	3	24.4	47.8	3921	2	US-08-567-375-3	Sequence 3, Appli
c	4	24.4	47.8	3921	2	US-08-587-680A-3	Sequence 3, Appli
c	5	24.4	47.8	5992	2	US-08-475-891A-3	Sequence 3, Appli
	6	24	47.1	1856	4	US-09-205-258-52	Sequence 52, Appl
	7	23.8	46.7	1011	4	US-09-976-594-1069	Sequence 1069, Ap
c	8	23.8	46.7	1312	4	US-09-976-594-886	Sequence 886, App
	9	23.8	46.7	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	10	23.4	45.9	6609	4	US-09-976-594-690	Sequence 690, App
c	11	23.4	45.9	786431	4	US-09-751-389-3	Sequence 3, Appli
	12	23.2	45.5	963	4	US-09-328-352-1552	Sequence 1552, Ap
	13	23.2	45.5	1506	4	US-09-134-001C-1278	Sequence 1278, Ap
c	14	23	45.1	1026	4	US-09-540-236-1082	Sequence 1082, Ap
c	15	23	45.1	19988	4	US-09-596-002-10	Sequence 10, Appl
c	16	22.8	44.7	458	4	US-09-387-286-35	Sequence 35, Appl
	17	22.8	44.7	5532	4	US-08-956-171E-530	Sequence 530, App
c	18	22.8	44.7	6256	2	US-08-475-891A-1	Sequence 1, Appli
c	19	22.8	44.7	6256	2	US-08-567-375-1	Sequence 1, Appli
c	20	22.8	44.7	6256	2	US-08-587-680A-1	Sequence 1, Appli
	21	22.6	44.3	466	4	US-09-621-976-13701	Sequence 13701, A
	22	22.6	44.3	561	4	US-09-601-198-172	Sequence 172, App
	23	22.4	43.9	446	4	US-09-621-976-9662	Sequence 9662, Ap
	24	22.4	43.9	966	4	US-09-107-532A-1038	Sequence 1038, Ap
	25	22.4	43.9	1731	4	US-09-134-001C-1118	Sequence 1118, Ap
	26	22.4	43.9	3885	4	US-09-328-352-2188	Sequence 2188, Ap
	27	22.4	43.9	5895	4	US-08-956-171E-1	Sequence 1, Appli
c	28	22.4	43.9	36159	4	US-09-749-588-3	Sequence 3, Appli
	29	22.2	43.5	813	4	US-09-134-001C-2724	Sequence 2724, Ap
c	30	22.2	43.5	1245	4	US-09-543-681A-3265	Sequence 3265, Ap
	31	22.2	43.5	1281	4	US-09-134-000C-1305	Sequence 1305, Ap
c	32	22.2	43.5	1618	4	US-09-800-729-29	Sequence 29, Appl
	33	22.2	43.5	1626	4	US-09-358-383C-6	Sequence 6, Appli
	34	22.2	43.5	2694	4	US-09-358-383C-4	Sequence 4, Appli
	35	22.2	43.5	3321	4	US-09-358-383C-17	Sequence 17, Appl
	36	22.2	43.5	5107	4	US-09-358-383C-15	Sequence 15, Appl
	37	22.2	43.5	5955	4	US-09-358-383C-14	Sequence 14, Appl
	38	22.2	43.5	202001	4	US-09-734-674-3	Sequence 3, Appli
	39	22	43.1	400	4	US-08-956-171E-2161	Sequence 2161, Ap
c	40	22	43.1	904	4	US-09-171-209-59	Sequence 59, Appl
c	41	22	43.1	1044	4	US-09-221-017B-113	Sequence 113, App
	42	22	43.1	1137	4	US-09-134-001C-657	Sequence 657, App
	43	22	43.1	1782	2	US-08-714-168-3	Sequence 3, Appli
	44	22	43.1	1782	3	US-09-320-721A-3	Sequence 3, Appli
c	45	22	43.1	5238	4	US-08-961-527-150	Sequence 150, App

ALIGNMENTS

RESULT 1
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragme
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Query Match 56.1%; Score 28.6; DB 4; Length 580073;
Best Local Similarity 72.5%; Pred. No. 1.4;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              || ||||| | ||||| | ||||| |||||||||| | ||| ||   | ||
Db     508476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 508526
```

RESULT 2
US-09-205-258-15
; Sequence 15, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881

```
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 1018
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-205-258-15

  Query Match             48.2%;  Score 24.6;  DB 4;  Length 1018;
  Best Local Similarity   70.2%;  Pred. No. 15;
  Matches   33;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATC 49
              ||| ||| |||| | | ||| |||||| |||||   | ||  |||
Db        746 AAATAGATATTGCTCATTAAGGTAAATATTTTTGTTGATGAATGATC 792


RESULT 3
US-08-567-375-3/c
; Sequence 3, Application US/08567375
; Patent No. 5952485
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Szabo, Veronique
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/567,375
;      FILING DATE:  04-DEC-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,645
;      FILING DATE:  29-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/475,891
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  023070-058930
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3921 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(1..2676, 3520..3918)
;      OTHER INFORMATION:  /product= "Xa-21"
US-08-567-375-3
```

```
  Query Match             47.8%;  Score 24.4;  DB 2;  Length 3921;
  Best Local Similarity   68.0%;  Pred. No. 19;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3456 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3407


RESULT 4
US-08-587-680A-3/c
; Sequence 3, Application US/08587680A
; Patent No. 5977434
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Szabo, Veronique
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/587,680A
;      FILING DATE:  17-JAN-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/475,891
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/004,645
;      FILING DATE:  29-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/567,375
;      FILING DATE:  04-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02307O-058940US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3921 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(1..2676, 3520..3918)
;      OTHER INFORMATION:  /product= "Xa-21"
US-08-587-680A-3

  Query Match             47.8%;  Score 24.4;  DB 2;  Length 3921;
  Best Local Similarity   68.0%;  Pred. No. 19;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3456 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3407


RESULT 5
US-08-475-891A-3/c
; Sequence 3, Application US/08475891A
; Patent No. 5859339
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resistance in Plants
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,891A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,375
;      FILING DATE:  17-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02370-058910US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5992 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(512..3149, 3993..4393)
;      OTHER INFORMATION:  /product= "RRK-B"
;      OTHER INFORMATION:  /note= "Xa21 Xanthomonas spp. disease
;      OTHER INFORMATION:  resistance gene RRK-B from rice (Oryza
;      OTHER INFORMATION:  sativa)"
US-08-475-891A-3

  Query Match             47.8%;  Score 24.4;  DB 2;  Length 5992;
  Best Local Similarity   68.0%;  Pred. No. 20;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||||||||| |||||  || | |||  ||  |||  |  | ||| ||||
Db       3928 ATAAAAAGAACTTGGATGTATAATTATGTAAATTTACGTGTTAAATATCA 3879


RESULT 6
US-09-205-258-52
; Sequence 52, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
```

```
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 1856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-205-258-52

  Query Match             47.1%;  Score 24;  DB 4;  Length 1856;
  Best Local Similarity   68.8%;  Pred. No. 24;
  Matches   33;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||| || ||  | |||||      |||| | | ||||||||| ||
Db        205 AAAAGTAAATGTCCATAAAGCGTTTCACTTATATTCTTCAAACATGAT 252


RESULT 7
US-09-976-594-1069
; Sequence 1069, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 1069
;   LENGTH: 1011
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 107309.1
;   NAME/KEY: unsure
;   LOCATION: 433-436, 445, 447, 454, 456, 463-465, 472, 495, 498, 662, 939
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-1069

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1011;
  Best Local Similarity   72.1%;  Pred. No. 27;
  Matches   31;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              |||||||| |  ||| ||||| || |  | |  ||||||| ||
Db        258 CAAAAAGATAGAGGATTTAAATTTCACAATTGATGTGCTTTAA 300


RESULT 8
US-09-976-594-886/c
; Sequence 886, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
```

```
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 886
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 981037.1
US-09-976-594-886

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1312;
  Best Local Similarity   80.0%;  Pred. No. 27;
  Matches   28;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGT 37
              ||| ||||||| || || |||||||||| |||| |
Db       1218 AAAGAGAAATTTTACATATAGTTAAATAATTTTTT 1184


RESULT 9
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (674435)..(674435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (741684)..(741684)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             46.7%;  Score 23.8;  DB 4;  Length 1664976;
  Best Local Similarity   66.7%;  Pred. No. 50;
  Matches   34;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              | ||| | |||||   | |  |  || ||| |||  ||| ||||||||  ||
Db     555193 AAAAAGAAAAATTAATCATGCAAATATATAATTTAATGCATCAAACATTAT 555243


RESULT 10
US-09-976-594-690/c
; Sequence 690, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 690
;   LENGTH: 6609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 175918.15
;   NAME/KEY: unsure
;   LOCATION: 825
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-690

  Query Match             45.9%;  Score 23.4;  DB 4;  Length 6609;
  Best Local Similarity   67.3%;  Pred. No. 43;
  Matches   33;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATC 49
              || |||| ||||||     |||| ||||    |||   |||||||| |
Db       5196 ACCAAAATAAATTGTGAAAAAAGATAAAAGACTTTCATCTTCAAACAAC 5148


RESULT 11
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001067
;  CURRENT APPLICATION NUMBER: US/09/751,389
;  CURRENT FILING DATE:  2001-01-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 786431
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(786431)
;   OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

  Query Match             45.9%;  Score 23.4;  DB 4;  Length 786431;
  Best Local Similarity   81.8%;  Pred. No. 65;
  Matches   27;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         18 TTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||| ||||| |||||||||| ||||| | ||
Db     658324 TGAAACTTAAAAACTTTTGTGCCTCAAAAAACA 658292


RESULT 12
US-09-328-352-1552
; Sequence 1552, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
```

```
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1552
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-1552

  Query Match             45.5%;  Score 23.2;  DB 4;  Length 963;
  Best Local Similarity   70.5%;  Pred. No. 42;
  Matches   31;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAA 44
              | |  ||  || ||||||||||  ||  | |||||| ||| ||
Db        263 ATATTAATGAAGTGGACTTAAATGTACGTCCTTTTGTTCTTAAA 306


RESULT 13
US-09-134-001C-1278
; Sequence 1278, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1278
;   LENGTH: 1506
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1278

  Query Match             45.5%;  Score 23.2;  DB 4;  Length 1506;
  Best Local Similarity   77.8%;  Pred. No. 43;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 CAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGT 37
              |||||||||| |||  |||||| |||| | || | |
Db       1360 CAAAAAGAAACTGGTATTAAAGGTAAACAATTATTT 1395


RESULT 14
US-09-540-236-1082/c
; Sequence 1082, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
;  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.2005-001
;  CURRENT APPLICATION NUMBER: US/09/540,236
;  CURRENT FILING DATE: 2000-04-04
;  NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1082
;   LENGTH: 1026
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-1082

  Query Match             45.1%;  Score 23;  DB 4;  Length 1026;
  Best Local Similarity   68.1%;  Pred. No. 49;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||   |||  ||  ||||| |||| | ||||  ||| | ||| ||||
Db        394 AATCCAAAAAGGCTTTAAATTTAATTTCTTTGTTGCATAAAATATCA 348


RESULT 15
US-09-596-002-10/c
; Sequence 10, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
;  APPLICANT:      Lagace, Robert, E.
;  APPLICANT:  Patterson, Chandra
;  APPLICANT:  Berg, Kim, L.
;  TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
;  FILE REFERENCE:  PM-0008-4 US
;  CURRENT APPLICATION NUMBER: US/09/596,002
;  CURRENT FILING DATE:  2000-06-16
;  PRIOR APPLICATION NUMBER:  60/140,121
;  PRIOR FILING DATE:  1999-06-18
;  NUMBER OF SEQ ID NOS:      41
;  SOFTWARE:      PERL Program
; SEQ ID NO 10
;   LENGTH: 19988
;   TYPE: DNA
;   ORGANISM: M. catarrhalis
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte template ID No. 6632636 10
;  PUBLICATION INFORMATION:
US-09-596-002-10

  Query Match             45.1%;  Score 23;  DB 4;  Length 19988;
  Best Local Similarity   68.1%;  Pred. No. 65;
  Matches   32;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||   |||  ||  ||||| |||| | ||||  ||| | ||| ||||
Db       1287 AATCCAAAAAGGCTTTAAATTTAATTTCTTTGTTGCATAAAATATCA 1241


Search completed: May  7, 2004, 15:44:33
Job time : 34.6094 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 11:56:28 ; Search time 104.189 Seconds
(without alignments)
2079.475 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747720

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	48	94.1	168174	6	ABT11173	Abt11173 Human 5-l
	2	48	94.1	168273	6	ABT11114	Abt11114 Human 5-l
	3	38.4	75.3	13249	6	ABL32116	Abl32116 Human imm
c	4	38.4	75.3	13249	6	ABL32117	Abl32117 Human imm
c	5	38.4	75.3	13249	6	ABK31177	Abk31177 Signal tr
	6	38.4	75.3	13249	6	ABK31176	Abk31176 Signal tr
c	7	38.4	75.3	13249	6	ABL70132	Abl70132 Chemicall
	8	38.4	75.3	13249	6	ABL70131	Abl70131 Chemicall
c	9	28.6	56.1	879	7	ACA39547	Aca39547 Prokaryot
c	10	28.6	56.1	10809	7	ACC69139	Acc69139 M. genita
	11	28.6	56.1	80073	2	AAT58840_5	Continuation (6 of
	12	28.6	56.1	110000	2	AAT58840_4	Continuation (5 of
c	13	26.4	51.8	355	4	AAI83146	Aai83146 Human pol
c	14	25.6	50.2	5864	6	ABL33441	Abl33441 Human imm
c	15	25.6	50.2	5864	6	ABL54362	Abl54362 Chemicall
c	16	25.6	50.2	349938	9	ADC87621	Adc87621 Human GPC
	17	25.4	49.8	607	5	ABV58521	Abv58521 Human pro
	18	25.2	49.4	206	4	AAK81112	Aak81112 Human imm
c	19	25.2	49.4	1214	3	AAZ93368	Aaz93368 Sequence
c	20	25.2	49.4	1214	6	AAL41059	Aal41059 cDNA of H
c	21	25.2	49.4	1285	4	ABK41870	Abk41870 cDNA enco
c	22	25.2	49.4	1285	8	ADB59537	Adb59537 Connectiv
	23	25.2	49.4	1779	3	AAC80574	Aac80574 Human sec
	24	25.2	49.4	5728	4	AAK81109	Aak81109 Human imm
	25	25.2	49.4	5733	4	AAK81111	Aak81111 Human imm
	26	25.2	49.4	5733	4	AAK81110	Aak81110 Human imm
c	27	25.2	49.4	18434	6	ABL34006	Abl34006 Human imm
c	28	25	49.0	4129	4	ABL17308	Abl17308 Drosophil
c	29	25	49.0	9888	6	ABL33240	Abl33240 Human imm
	30	25	49.0	110000	7	AAL52246_1	Continuation (2 of
c	31	25	49.0	240000	7	ACD13446	Acd13446 Human DNA
	32	24.6	48.2	446	4	AAI81019	Aai81019 Human pol
	33	24.6	48.2	948	6	ABK72766	Abk72766 Bacillus
	34	24.6	48.2	1018	2	AAV84415	Aav84415 Human sec
	35	24.6	48.2	1018	4	ABA83198	Aba83198 Human sec
	36	24.6	48.2	1018	8	ACH04699	Ach04699 Novel hum
	37	24.6	48.2	1018	8	ACD44509	Acd44509 Human cDN
	38	24.6	48.2	1218	6	ABK73006	Abk73006 Bacillus
	39	24.6	48.2	3982	4	AAK78949	Aak78949 Human imm
	40	24.6	48.2	4460	6	AAD48267	Aad48267 Ehrlichia
c	41	24.6	48.2	5678	6	ABL33138	Abl33138 Human imm
	42	24.6	48.2	22927	4	AAL04782	Aal04782 Human rep
	43	24.6	48.2	22927	4	ABL97677	Abl97677 Human tes
	44	24.6	48.2	87878	8	ADA02576	Ada02576 Human FKB
	45	24.6	48.2	87878	9	ADB72314	Adb72314 Human FKB

ALIGNMENTS

RESULT 1
ABT11173
ID ABT11173 standard; DNA; 168174 BP.
XX
AC ABT11173;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 63.
XX
KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200262825-A2.
XX
PD 15-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-US003546.
XX
PR 08-FEB-2001; 2001US-0267515P.
PR 21-AUG-2001; 2001US-0314248P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes G, Meyer J;
XX
DR WPI; 2002-627522/67.
XX
PT New isolated nucleic acid molecule with an allelic variant of a
PT polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT prognosticating disorders associated with an aberrant inflammatory
PT response such as asthma.
XX
PS Disclosure; Fig 4; 290pp; English.
XX
CC The invention relates to an isolated human nucleic acid molecule
CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC selected from any of 3, 20 or 21 base pair sequences, given in the
CC specification, or their complement. The compositions and methods of the

CC present invention are useful for diagnosing and/or prognosing disorders
CC associated with an aberrant inflammatory response such as asthma,
CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC disease, atopic and contact dermatitis. The nucleic acid molecules can
CC also be useful for identifying an individual amongst other individuals
CC from the same species for use in forensic medicine and paternity testing.
CC This polynucleotide sequence represents DNA relating to the human 5-
CC lipoxygenase (5-LO) gene of the invention
XX
SQ Sequence 168174 BP; 46808 A; 36442 C; 36942 G; 46474 T; 0 U; 1508 Other;

Query Match 94.1%; Score 48; DB 6; Length 168174;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 166822 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 166869

RESULT 2
ABT11114
ID ABT11114 standard; DNA; 168273 BP.
XX
AC ABT11114;
XX
DT 05-DEC-2002 (first entry)
XX
DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 2.
XX
KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200262825-A2.
XX
PD 15-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-US003546.
XX
PR 08-FEB-2001; 2001US-0267515P.
PR 21-AUG-2001; 2001US-0314248P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes G, Meyer J;
XX
DR WPI; 2002-627522/67.
XX
PT New isolated nucleic acid molecule with an allelic variant of a
PT polymorphic region of an 5-LO gene, useful for diagnosing and/or
PT prognosticating disorders associated with an aberrant inflammatory
PT response such as asthma.
XX
PS Disclosure; Fig 2; 290pp; English.
XX
CC The invention relates to an isolated human nucleic acid molecule
CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
CC (5-LO) gene, where the allelic variant comprises one or more nucleotide
CC selected from any of 3, 20 or 21 base pair sequences, given in the
CC specification, or their complement. The compositions and methods of the
CC present invention are useful for diagnosing and/or prognosing disorders
CC associated with an aberrant inflammatory response such as asthma,
CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC disease, atopic and contact dermatitis. The nucleic acid molecules can
CC also be useful for identifying an individual amongst other individuals
CC from the same species for use in forensic medicine and paternity testing.
CC This polynucleotide sequence represents DNA relating to the human 5-
CC lipoxygenase (5-LO) gene of the invention
XX
SQ Sequence 168273 BP; 46834 A; 36467 C; 36966 G; 46498 T; 0 U; 1508 Other;

Query Match 94.1%; Score 48; DB 6; Length 168273;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 166921 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 166968

RESULT 3
ABL32116
ID ABL32116 standard; DNA; 13249 BP.
XX
AC ABL32116;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 89.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 89; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

Query Match 75.3%; Score 38.4; DB 6; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.0082;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709

RESULT 4
ABL32117/c
ID   ABL32117 standard; DNA; 13249 BP.
XX
AC   ABL32117;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 90.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 90; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541

RESULT 5
ABK31177/c
ID   ABK31177 standard; DNA; 13249 BP.
XX
AC   ABK31177;
XX
DT   23-APR-2002  (first entry)
XX
DE   Signal transduction associated gene modified complementary DNA #10.
XX
KW   Human; signal transduction associated gene; cytosine methylation state;
KW   CpG island; signal transduction associated disease; solid tumour; cancer;
KW   antitumour; cytostatic; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007472.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction.
XX
PS   Claim 1; SEQ ID NO 20; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC   Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC   the cytosine methylation state (CpG islands) of these genes, and a method
CC   for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC   genes associated with signal transduction. The genomic DNA can be
CC   obtained from cells or cellular components which contain DNA, e.g. cell
CC   lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC   tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC   brain, heart, prostate, lung, breast or liver, histologic object slides,
CC   and all their possible combinations. The sequences of the invention are
CC   useful for the diagnosis and therapy of diseases associated with signal
CC   transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC   chemically pretreated genomic DNA sequences of different genes associated
CC   with signal transduction, or their complementary sequences. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541

RESULT 6
ABK31176
ID   ABK31176 standard; DNA; 13249 BP.
XX
AC   ABK31176;
XX
DT   23-APR-2002  (first entry)
XX
DE   Signal transduction associated gene modified DNA #10.
```

```
XX
KW   Human; signal transduction associated gene; cytosine methylation state;
KW   CpG island; signal transduction associated disease; solid tumour; cancer;
KW   antitumour; cytostatic; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007472.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction.
XX
PS   Claim 1; SEQ ID NO 19; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC   Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC   the cytosine methylation state (CpG islands) of these genes, and a method
CC   for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC   genes associated with signal transduction. The genomic DNA can be
CC   obtained from cells or cellular components which contain DNA, e.g. cell
CC   lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC   tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC   brain, heart, prostate, lung, breast or liver, histologic object slides,
CC   and all their possible combinations. The sequences of the invention are
CC   useful for the diagnosis and therapy of diseases associated with signal
CC   transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC   chemically pretreated genomic DNA sequences of different genes associated
CC   with signal transduction, or their complementary sequences. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 7
ABL70132/c
ID   ABL70132 standard; DNA; 13249 BP.
XX
AC   ABL70132;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated cell signalling DNA sequence complementary to#11.
XX
KW   Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW   tumour; cytostatic; ds.
XX
OS   Unidentified.
XX
PN   WO200202807-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007471.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154758/20.
XX
PT   Nucleic acid, useful for diagnosis and therapy of diseases associated
PT   with cell signaling e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with cell signaling.
XX
PS   Claim 1; SEQ ID NO 22; 24pp + Sequence Listing; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of at least
CC   18 bases of a segment of chemically pretreated DNA of genes associated
CC   with cell signalling. The activity of the modified sequences of the
CC   invention may be described as cytostatic. The object of the invention is
CC   to provide the chemically modified DNA of genes associated with cell
CC   signalling, as well as oligonucleotides and/or PNA-oligomers for
CC   detecting cytosine methylations, as well as a method which is
CC   particularly suitable for the diagnosis and/or therapy of genetic and
CC   epigenetic parameters of genes associated with cell signalling. The
CC   chemically modified DNA provided by the invention is useful for diagnosis
CC   and therapy of diseases such as solid tumours and cancer. The sequences
CC   given in records ABL70111-ABL70626 represent chemically pre-treated
CC   genomic DNA's of genes associated with cell signalling. Note: The
CC   sequence data for this patent is not represented in the printed
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office
XX
SQ   Sequence 13249 BP; 3128 A; 273 C; 3397 G; 6451 T; 0 U; 0 Other;

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.0082;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 8
ABL70131
ID   ABL70131 standard; DNA; 13249 BP.
XX
AC   ABL70131;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated cell signalling DNA sequence#11.
XX
KW   Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW   tumour; cytostatic; ds.
XX
OS   Unidentified.
XX
PN   WO200202807-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007471.
XX
PR   30-JUN-2000; 2000DE-01032529.
```

PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signaling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signaling.
XX
PS Claim 1; SEQ ID NO 21; 24pp + Sequence Listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office
XX
SQ Sequence 13249 BP; 3594 A; 273 C; 3130 G; 6252 T; 0 U; 0 Other;

Query Match 75.3%; Score 38.4; DB 6; Length 13249;
Best Local Similarity 87.5%; Pred. No. 0.0082;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
||||||||||||| |||||||||||||| ||||||| || ||| || ||
Db 3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709

RESULT 9
ACA39547/c
ID ACA39547 standard; DNA; 879 BP.
XX
AC ACA39547;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #21204.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycoplasma genitalium.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU35677.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 27417; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 879 BP; 269 A; 120 C; 143 G; 347 T; 0 U; 0 Other;

Query Match 56.1%; Score 28.6; DB 7; Length 879;
Best Local Similarity 72.5%; Pred. No. 8.8;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
|| ||||| | ||||| | ||||| |||||||||| | ||| || | ||
Db 537 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 487

RESULT 10
ACC69139/c
ID ACC69139 standard; DNA; 10809 BP.
XX
AC ACC69139;
XX
DT 10-JUL-2003 (first entry)
XX
DE M. genitalium aerobic metabolism gene cassette DNA SEQ ID NO:7.
XX
KW Mycoplasma genitalium; gene cassette; replication; transcription;
KW translation; metabolism; basic genetic operating system; gene therapy;
KW autonomous prototrophic nanomachine; auxotrophic nanomachine;
KW nanomachine; bioreactor; bioremediation; therapeutic; delivery system;
KW artificial tissue; artificial organ system; energy conversion system;
KW processing system; anabolic system; catabolic system; biological film;
KW biological coating; cosmetic; gene; ds.
XX
OS Mycoplasma genitalium.
XX

PN WO2003025145-A2.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-US029811.
XX
PR 20-SEP-2001; 2001US-00960870.
XX
PA (EGEA-) EGEA BIOSCIENCES INC.
XX
PI Evans GA;
XX
DR WPI; 2003-354602/33.
XX
PT New basic genetic operating system for autonomous prototrophic or
PT auxotrophic nanomachine, useful for therapeutic, diagnostic or industrial
PT purposes, comprises a nanomachine genome encoding a gene set for
PT viability or replication.
XX
PS Example 1; Page 210-213; 250pp; English.
XX
CC The present invention describes a basic genetic operating system for an
CC autonomous prototrophic or auxotrophic nanomachine comprising a
CC nanomachine genome encoding a minimal gene set sufficient for viability
CC or replication,.optionally in the presence of an auxotrophic molecule.
CC Also described is an autonomous prototrophic or auxotrophic nanomachine
CC comprising a basic genetic operating system for autonomous prototrophic
CC or auxotrophic viability or replication, optionally in the presence of an
CC auxotrophic molecule, and a particle envelope. The nanomachines can be
CC used in gene therapy. The basic genetic operating system or nanomachine
CC is useful in therapeutic, diagnostic and industrial applications, e.g. as
CC a bioreactor, for bioremediation, for the production of a therapeutic
CC biomolecule or as a therapeutic reagent, for the production of a
CC diagnostic indicator or reagent, as a delivery system, as an artificial
CC tissue or organ system, as an energy conversion system, as a processing
CC system, as an anabolic or catabolic system, for the production of
CC biological films or coatings, and for cosmetic applications. The present
CC sequence represents a Mycoplasma genitalium gene cassette nucleotide
CC sequence, which is used in an example from the present invention for the
CC design and synthesis of a basic genetic operation system for a
CC replication competent nanomachine
XX
SQ Sequence 10809 BP; 3805 A; 1468 C; 1932 G; 3604 T; 0 U; 0 Other;

Query Match 56.1%; Score 28.6; DB 7; Length 10809;
Best Local Similarity 72.5%; Pred. No. 10;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
|| ||||| | ||||| | ||||| ||||||||| | ||| || | ||
Db 4241 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 4191

RESULT 11
AAT58840_5
Continuation (6 of 6) of AAT58840 from base 500001 (Mycoplasma genitalium genome.)
WP Sequence split into 6 fragments LOCUS AAT58840 Accession Aat58840
WP Fragment Name Begin End
WP AAT58840_0 1 110000
WP AAT58840_1 100001 210000
WP AAT58840_2 200001 310000
WP AAT58840_3 300001 410000
WP AAT58840_4 400001 510000
WP AAT58840_5 500001 580073

Query Match 56.1%; Score 28.6; DB 2; Length 80073;
Best Local Similarity 72.5%; Pred. No. 12;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
|| ||||| | ||||| | ||||| ||||||||| | ||| || | ||
Db 8476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 8526

RESULT 12
AAT58840_4
Continuation (5 of 6) of AAT58840 from base 400001 (Mycoplasma genitalium genome.)
WP Sequence split into 6 fragments LOCUS AAT58840 Accession Aat58840
WP Fragment Name Begin End
WP AAT58840_0 1 110000
WP AAT58840_1 100001 210000
WP AAT58840_2 200001 310000
WP AAT58840_3 300001 410000
WP AAT58840_4 400001 510000
WP AAT58840_5 500001 580073

Query Match 56.1%; Score 28.6; DB 2; Length 110000;
Best Local Similarity 72.5%; Pred. No. 12;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
|| ||||| | ||||| | ||||| ||||||||| | ||| || | ||
Db 108476 ACTAAAAGGATTTGGAATGAAAGTAGAATACTTTTTTCCTTTAACAGTAAT 108526

RESULT 13
AAI83146/c
ID AAI83146 standard; cDNA; 355 BP.
XX
AC AAI83146;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 3206.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO03215.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 3206; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

```
XX
SQ   Sequence 355 BP; 86 A; 65 C; 74 G; 130 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 26.4;  DB 4;  Length 355;
  Best Local Similarity   75.0%;  Pred. No. 41;
  Matches   33;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 AGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | ||||||||||||||| || || ||| |||  ||| |||   ||
Db        209 ATAAATTGGACTTAAAATTCAAAACTCTTGCCCTTTAAAAGACA 166


RESULT 14
ABL33441/c
ID   ABL33441 standard; DNA; 5864 BP.
XX
AC   ABL33441;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1414.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 1414; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 5864 BP; 1601 A; 136 C; 1351 G; 2776 T; 0 U; 0 Other;

  Query Match             50.2%;  Score 25.6;  DB 6;  Length 5864;
  Best Local Similarity   70.8%;  Pred. No. 89;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              | ||||| ||||   || ||||   ||| |||| | | ||||||| ||
Db       4241 AAAAAAACAAATAAAACCTAAAAACAAAAACTTATTTTCTTCAAAAAT 4194


RESULT 15
ABL54362/c
ID   ABL54362 standard; DNA; 5864 BP.
XX
AC   ABL54362;
XX
DT   29-JUL-2002  (first entry)
XX
DE   Chemically treated apoptosis gene complementary to gene #31.
XX
KW   Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;
KW   Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;
KW   cancer; ds.
XX
OS   Unidentified.
XX
PN   WO200177164-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-EP003969.
XX
PR   06-APR-2000; 2000DE-01019058.
PR   07-APR-2000; 2000DE-01019173.
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-017444/02.
XX
PT   Chemically modified sequences of genes associated with apoptosis are
PT   useful to determine methylation patterns of genomic DNA samples for
PT   diagnosis of associated diseases such as cancer.
XX
PS   Claim 1; Seq ID #62; 24pp; English.
XX
CC   This invention relates to chemically pre-treated DNA of genes associated
CC   with apoptosis. The nucleic acids are used to allocate patients for
CC   specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,
CC   neurodegenerative disorders, Herpes simplex virus infection, renal
CC   ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This
CC   nucleotide sequence represents a chemically treated apoptosis gene. Even
CC   SEQ ID numbers are the complementary DNA strands to the odd SEQ ID
CC   numbers. The sequence data for this patent is not represented in the
CC   printed specification but is based on information supplied by the
CC   European patent office
XX
SQ   Sequence 5864 BP; 1601 A; 136 C; 1351 G; 2776 T; 0 U; 0 Other;

  Query Match             50.2%;  Score 25.6;  DB 6;  Length 5864;
  Best Local Similarity   70.8%;  Pred. No. 89;
  Matches   34;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACAT 48
              | ||||| ||||   || ||||   ||| |||| | | ||||||| ||
Db       4241 AAAAAAACAAATAAAACCTAAAAACAAAAACTTATTTTCTTCAAAAAT 4194


Search completed: May  7, 2004, 13:50:20
Job time : 111.189 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 13:24:59 ; Search time 531.669 Seconds
(without alignments)
4157.648 Million cell updates/sec

Title: US-10-071-411A-1_COPY_450_500
Perfect score: 51
Sequence: 1 acaaaaagaaattggactta.....ttttgtgcttcaaacatcat 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940541

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		48	94.1	129266	9	AL731567	AL731567 Human DNA
2		48	94.1	160654	2	AC011879	AC011879 Homo sapi
3		48	94.1	194453	2	AC010862	AC010862 Homo sapi
4		38.4	75.3	13249	6	AX344172	AX344172 Sequence
5	c	38.4	75.3	13249	6	AX344173	AX344173 Sequence
6		38.4	75.3	13249	6	AX345018	AX345018 Sequence
7	c	38.4	75.3	13249	6	AX345019	AX345019 Sequence
8		38.4	75.3	13249	6	AX348563	AX348563 Sequence
9	c	38.4	75.3	13249	6	AX348564	AX348564 Sequence
10	c	34	66.7	192044	9	AL590439	AL590439 Human DNA
11	c	33.6	65.9	128529	2	AC025758	AC025758 Homo sapi
12		33.6	65.9	157325	9	AC008810	AC008810 Homo sapi
13		33.6	65.9	164217	9	AC093264	AC093264 Homo sapi
14	c	33.4	65.5	209016	9	BS000239	BS000239 Pan trogl
15	c	32.6	63.9	173053	10	AL365334	AL365334 Mouse DNA
16	c	31.8	62.4	154157	9	HS101D08	AL133492 Homo sapi
17	c	31.8	62.4	223201	9	HS53I10	AL133493 Homo sapi
18	c	31.8	62.4	283388	2	AC012285	AC012285 Homo sapi
19	c	31.8	62.4	340000	9	HS21C103	AL163303 Homo sapi
20	c	30.6	60.0	60812	2	AC145991	AC145991 Pan trogl
21		30.6	60.0	82419	9	AC004979	AC004979 Homo sapi
22	c	30.6	60.0	133691	9	AC074347	AC074347 Homo sapi
23	c	30.2	59.2	58250	2	AC103690	AC103690 Homo sapi
24		30.2	59.2	144017	2	AF235106	AF235106 Homo sapi
25		30.2	59.2	169537	9	AC100814	AC100814 Homo sapi
26	c	30.2	59.2	175910	2	AC091004	AC091004 Homo sapi
27	c	30.2	59.2	181907	2	AC108731	AC108731 Homo sapi
28		29.8	58.4	160307	9	AC018359	AC018359 Homo sapi
29		29.8	58.4	171347	9	AC099776	AC099776 Homo sapi
30		29.4	57.6	174873	9	AC009069	AC009069 Homo sapi
31	c	29	56.9	83250	9	HS377F16	Z93783 Human DNA s
32		29	56.9	143878	2	AL359974	AL359974 Homo sapi
33	c	29	56.9	146250	2	AC074240	AC074240 Homo sapi
34		29	56.9	159231	9	AL161654	AL161654 Human DNA
35		29	56.9	160629	9	AC073326	AC073326 Homo sapi
36		29	56.9	168608	2	AL591477	AL591477 Homo sapi
37		29	56.9	179876	9	AC087863	AC087863 Homo sapi
38		29	56.9	191830	2	AC026332	AC026332 Homo sapi
39		28.8	56.5	202495	9	CNS01DW6	AL136418 Human chr
40		28.8	56.5	202496	9	CNS01DX6	AL139054 Human chr
41		28.6	56.1	9374	1	U39722	U39722 Mycoplasma
42		28.6	56.1	80073	6	AR300198_5	Continuation (6 of
43		28.6	56.1	87790	9	AC090698	AC090698 Homo sapi
44		28.6	56.1	110000	6	AR300198_4	Continuation (5 of
45	c	28.6	56.1	161198	2	AC015867	AC015867 Homo sapi

ALIGNMENTS

RESULT 1
AL731567
LOCUS AL731567 129266 bp DNA linear PRI 20-JUN-2002
DEFINITION Human DNA sequence from clone RP11-67C2 on chromosome 10, complete sequence.
ACCESSION AL731567 AC010865
VERSION AL731567.6 GI:21537524
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 129266)
AUTHORS Whitehead,S.
TITLE Direct Submission

```
  JOURNAL   Submitted (31-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 21, 2002 this sequence version replaced gi:21213582.
            Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver
            Street, Waltham, MA 02453, USA
            http://www.genomecorp.com
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-67C2 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .129266
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-67C2"
                     /clone_lib="RPCI-11.1"
ORIGIN

  Query Match             94.1%;  Score 48;  DB 9;  Length 129266;
  Best Local Similarity   100.0%;  Pred. No. 0.00039;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33190 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 33237


RESULT 2
AC011879
LOCUS       AC011879               160654 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-16P14, WORKING DRAFT SEQUENCE, 30 unordered
            pieces.
ACCESSION   AC011879
VERSION     AC011879.3  GI:7239554
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-16P14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160654)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 14, 2000 this sequence version replaced gi:6524208.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3606
                Center clone name: 16_P_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 111055 bases at least Q40
                Consensus quality: 135066 bases at least Q30
                Consensus quality: 147921 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 157754; sum-of-contigs
                Quality coverage: 2.9 in Q20 bases; agarose-fp
                Quality coverage: 3.0 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 30 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      151: contig of 151 bp in length
            *      152      251: gap of 100 bp
            *      252     1760: contig of 1509 bp in length
            *     1761     1860: gap of 100 bp
            *     1861     3069: contig of 1209 bp in length
            *     3070     3169: gap of 100 bp
            *     3170     4720: contig of 1551 bp in length
            *     4721     4820: gap of 100 bp
            *     4821     6174: contig of 1354 bp in length
            *     6175     6274: gap of 100 bp
            *     6275     7417: contig of 1143 bp in length
            *     7418     7517: gap of 100 bp
            *     7518     9158: contig of 1641 bp in length
            *     9159     9258: gap of 100 bp
            *     9259    10865: contig of 1607 bp in length
            *    10866    10965: gap of 100 bp
            *    10966    12859: contig of 1894 bp in length
            *    12860    12959: gap of 100 bp
            *    12960    15671: contig of 2712 bp in length
            *    15672    15771: gap of 100 bp
            *    15772    18082: contig of 2311 bp in length
            *    18083    18182: gap of 100 bp
            *    18183    20523: contig of 2341 bp in length
            *    20524    20623: gap of 100 bp
            *    20624    22903: contig of 2280 bp in length
            *    22904    23003: gap of 100 bp
            *    23004    23671: contig of 668 bp in length
            *    23672    23771: gap of 100 bp
```

```
            *    23772    25541: contig of 1770 bp in length
            *    25542    25641: gap of 100 bp
            *    25642    28323: contig of 2682 bp in length
            *    28324    28423: gap of 100 bp
            *    28424    31498: contig of 3075 bp in length
            *    31499    31598: gap of 100 bp
            *    31599    36626: contig of 5028 bp in length
            *    36627    36726: gap of 100 bp
            *    36727    42109: contig of 5383 bp in length
            *    42110    42209: gap of 100 bp
            *    42210    48339: contig of 6130 bp in length
            *    48340    48439: gap of 100 bp
            *    48440    55333: contig of 6894 bp in length
            *    55334    55433: gap of 100 bp
            *    55434    63594: contig of 8161 bp in length
            *    63595    63694: gap of 100 bp
            *    63695    73943: contig of 10249 bp in length
            *    73944    74043: gap of 100 bp
            *    74044    83665: contig of 9622 bp in length
            *    83666    83765: gap of 100 bp
            *    83766    95322: contig of 11557 bp in length
            *    95323    95422: gap of 100 bp
            *    95423   108403: contig of 12981 bp in length
            *   108404   108503: gap of 100 bp
            *   108504   120220: contig of 11717 bp in length
            *   120221   120320: gap of 100 bp
            *   120321   132958: contig of 12638 bp in length
            *   132959   133058: gap of 100 bp
            *   133059   145697: contig of 12639 bp in length
            *   145698   145797: gap of 100 bp
            *   145798   160654: contig of 14857 bp in length.
FEATURES             Location/Qualifiers
     source          1. .160654
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-16P14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .151
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    252. .1760
                     /note="assembly_fragment"
     misc_feature    1861. .3069
                     /note="assembly_fragment"
     misc_feature    3170. .4720
                     /note="assembly_fragment"
     misc_feature    4821. .6174
                     /note="assembly_fragment"
     misc_feature    6275. .7417
                     /note="assembly_fragment"
     misc_feature    7518. .9158
                     /note="assembly_fragment"
     misc_feature    9259. .10865
                     /note="assembly_fragment"
     misc_feature    10966. .12859
                     /note="assembly_fragment"
     misc_feature    12960. .15671
                     /note="assembly_fragment"
     misc_feature    15772. .18082
                     /note="assembly_fragment"
     misc_feature    18183. .20523
                     /note="assembly_fragment"
     misc_feature    20624. .22903
                     /note="assembly_fragment"
     misc_feature    23004. .23671
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    23772. .25541
                     /note="assembly_fragment"
     misc_feature    25642. .28323
                     /note="assembly_fragment"
     misc_feature    28424. .31498
                     /note="assembly_fragment"
     misc_feature    31599. .36626
                     /note="assembly_fragment"
     misc_feature    36727. .42109
                     /note="assembly_fragment"
     misc_feature    42210. .48339
                     /note="assembly_fragment"
     misc_feature    48440. .55333
                     /note="assembly_fragment"
     misc_feature    55434. .63594
                     /note="assembly_fragment"
     misc_feature    63695. .73943
                     /note="assembly_fragment"
     misc_feature    74044. .83665
                     /note="assembly_fragment"
     misc_feature    83766. .95322
                     /note="assembly_fragment"
     misc_feature    95423. .108403
                     /note="assembly_fragment"
     misc_feature    108504. .120220
                     /note="assembly_fragment"
     misc_feature    120321. .132958
                     /note="assembly_fragment"
     misc_feature    133059. .145697
                     /note="assembly_fragment"
     misc_feature    145798. .160654
                     /note="assembly_fragment"
ORIGIN

  Query Match             94.1%;  Score 48;  DB 2;  Length 160654;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40267 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 40314


RESULT 3
AC010862
LOCUS       AC010862              194453 bp    DNA     linear   HTG 30-AUG-2001
DEFINITION  Homo sapiens chromosome 06 clone RP11-326D18, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC010862
VERSION     AC010862.7  GI:9957987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 194453)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 194453)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Sep 1, 2000 this sequence version replaced gi:8247773.
            -------------- Genome Center
                           Center: Genome Therapeutics Corperation
                           Center code: GTC
                           Web site: http://www.genomecorp.com/
                           Contact: gtc-seqcenter@genomecorp.com
                        -------------- Project Information
                           Center project name: hg024
                        -------------- Summary Statistics
                           Sequencing vector: N/A
                           Chemistry: Dye-terminator Big Dye; 100% of reads
```

```
                     Assembly program: Phrap; version 990315
                     Consensus quality: 162991 bases at least Q40
                     Consensus quality: 176452 bases at least Q30
                     Consensus quality: 179870 bases at least Q20
                     Insert size: 192053; sum-of-contigs
                     Quality coverage: 4.2x in Q20 bases; sum-of-contigs
                   --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1117: contig of 1117 bp in length
            *    1118     1217: gap of unknown length
            *    1218     2365: contig of 1148 bp in length
            *    2366     2465: gap of unknown length
            *    2466     3576: contig of 1111 bp in length
            *    3577     3676: gap of unknown length
            *    3677     4782: contig of 1106 bp in length
            *    4783     4882: gap of unknown length
            *    4883     6049: contig of 1167 bp in length
            *    6050     6149: gap of unknown length
            *    6150     7670: contig of 1521 bp in length
            *    7671     7770: gap of unknown length
            *    7771     9521: contig of 1751 bp in length
            *    9522     9621: gap of unknown length
            *    9622    10895: contig of 1274 bp in length
            *   10896    10995: gap of unknown length
            *   10996    12183: contig of 1188 bp in length
            *   12184    12283: gap of unknown length
            *   12284    13602: contig of 1319 bp in length
            *   13603    13702: gap of unknown length
            *   13703    15955: contig of 2253 bp in length
            *   15956    16055: gap of unknown length
            *   16056    18297: contig of 2242 bp in length
            *   18298    18397: gap of unknown length
            *   18398    24368: contig of 5971 bp in length
            *   24369    24468: gap of unknown length
            *   24469    30049: contig of 5581 bp in length
            *   30050    30149: gap of unknown length
            *   30150    36961: contig of 6812 bp in length
            *   36962    37061: gap of unknown length
            *   37062    43997: contig of 6936 bp in length
            *   43998    44097: gap of unknown length
            *   44098    49940: contig of 5843 bp in length
            *   49941    50040: gap of unknown length
            *   50041    56987: contig of 6947 bp in length
            *   56988    57087: gap of unknown length
            *   57088    65541: contig of 8454 bp in length
            *   65542    65641: gap of unknown length
            *   65642    75225: contig of 9584 bp in length
            *   75226    75325: gap of unknown length
            *   75326    85420: contig of 10095 bp in length
            *   85421    85520: gap of unknown length
            *   85521   101885: contig of 16365 bp in length
            *  101886   101985: gap of unknown length
            *  101986   124008: contig of 22023 bp in length
            *  124009   124108: gap of unknown length
            *  124109   157199: contig of 33091 bp in length
            *  157200   157299: gap of unknown length
            *  157300   194453: contig of 37154 bp in length.
FEATURES             Location/Qualifiers
     source          1. .194453
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="06"
                     /clone="RP11-326D18"
                     /clone_lib="RPCI-11"

ORIGIN
```

```
  Query Match             94.1%;  Score 48;  DB 2;  Length 194453;
  Best Local Similarity   100.0%;  Pred. No. 0.00035;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4527 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 4574


RESULT 4
AX344172
LOCUS       AX344172               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 19 from Patent WO0200926.
ACCESSION   AX344172
VERSION     AX344172.1  GI:18492060
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 19 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 5
AX344173/c
LOCUS       AX344173               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 20 from Patent WO0200926.
ACCESSION   AX344173
VERSION     AX344173.1  GI:18492061
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with signal transduction
  JOURNAL   Patent: WO 0200926-A 20 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541
```

```
RESULT 6
AX345018
LOCUS       AX345018               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 89 from Patent WO0200928.
ACCESSION   AX345018
VERSION     AX345018.1  GI:18492904
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 89 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 7
AX345019/c
LOCUS       AX345019               13249 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 90 from Patent WO0200928.
ACCESSION   AX345019
VERSION     AX345019.1  GI:18492905
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 90 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 8
AX348563
LOCUS       AX348563               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 21 from Patent WO0202807.
ACCESSION   AX348563
VERSION     AX348563.1  GI:18614598
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 21 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||||||||||||| |||||||||||||| ||||||| || ||| || ||
Db       3662 AAAAGAAATTGGATTTAAAGTTAAATATTTTTGTGTTTTAAATATTAT 3709


RESULT 9
AX348564/c
LOCUS       AX348564               13249 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 22 from Patent WO0202807.
ACCESSION   AX348564
VERSION     AX348564.1  GI:18614599
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with cell signalling
  JOURNAL   Patent: WO 0202807-A 22 10-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .13249
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

  Query Match             75.3%;  Score 38.4;  DB 6;  Length 13249;
  Best Local Similarity   87.5%;  Pred. No. 0.33;
  Matches   42;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              |||| |||||  ||||||| |||||||||||| | ||||||||||||
Db       9588 AAAAAAAATTAAACTTAAAATTAAATACTTTTATACTTCAAACATCAT 9541


RESULT 10
AL590439/c
LOCUS       AL590439               192044 bp    DNA     linear   PRI 23-AUG-2001
DEFINITION  Human DNA sequence from clone RP11-394I23 on chromosome 10,
            complete sequence.
ACCESSION   AL590439
VERSION     AL590439.12  GI:15384822
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 192044)
  AUTHORS   Babbage,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
```

```
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 31, 2001 this sequence version replaced gi:14268248.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-394I23 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-394I23 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true right end of clone RP11-394I23 is at 192044 in this
            sequence. The true left end of clone RP11-657A9 is at 85254 in this
            sequence. The true right end of clone RP11-39J5 is at 100 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .192044
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-394I23"
                     /clone_lib="RPCI-11.2"
ORIGIN

  Query Match             66.7%;  Score 34;  DB 9;  Length 192044;
  Best Local Similarity   80.0%;  Pred. No. 2.6;
  Matches   40;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              ||||| | |||||||||||||| ||||| |||| |||| |||||   ||
Db      30825 ACAAAGATAAATTGGACTTAAAATTAAAAACTTGTGTGTGTCAAAGGACA 30776


RESULT 11
AC025758/c
LOCUS       AC025758              128529 bp    DNA     linear   HTG 18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone CTD-2235A13, WORKING DRAFT
            SEQUENCE, 16 ordered pieces.
ACCESSION   AC025758
VERSION     AC025758.3  GI:9256494
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 128529)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 128529)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:7711854.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 717802
            Center clone name: CITB-H1_2235A13
            -------
            Summary Statistics
            Consensus quality: 116698 bases at least Q40
            Consensus quality: 124565 bases at least Q30
            Consensus quality: 125980 bases at least Q20
            Estimated insert size: 130000; pulse field gel estimation
            Estimated insert size: 127829; sum-of-contigs estimation
            Quality coverage:  4.18 in Q20 bases; pulse field gel estimation
            Quality coverage:  4.25 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     8719: contig of 8719 bp in length
            *     8720     8819: gap of unknown length
            *     8820    17381: contig of 8562 bp in length
            *    17382    17481: gap of unknown length
            *    17482    22126: contig of 4645 bp in length
            *    22127    22226: gap of unknown length
            *    22227    25476: contig of 3250 bp in length
            *    25477    25576: gap of unknown length
            *    25577    36711: contig of 11135 bp in length
            *    36712    36811: gap of unknown length
            *    36812    52089: contig of 15278 bp in length
            *    52090    52189: gap of unknown length
            *    52190    56295: contig of 4106 bp in length
            *    56296    56395: gap of unknown length
            *    56396    59531: contig of 3136 bp in length
            *    59532    59631: gap of unknown length
            *    59632    66671: contig of 7040 bp in length
            *    66672    66771: gap of unknown length
            *    66772    83163: contig of 16392 bp in length
            *    83164    83263: gap of unknown length
            *    83264   100033: contig of 16770 bp in length
            *   100034   100133: gap of unknown length
            *   100134   102212: contig of 2079 bp in length
            *   102213   102312: gap of unknown length
            *   102313   104754: contig of 2442 bp in length
            *   104755   104854: gap of unknown length
            *   104855   112328: contig of 7474 bp in length
            *   112329   112428: gap of unknown length
            *   112429   124204: contig of 11776 bp in length
            *   124205   124304: gap of unknown length
            *   124305   128529: contig of 4225 bp in length.
FEATURES             Location/Qualifiers
     source          1. .128529
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2235A13"
                     /clone_lib="CalTech human BAC library D"
ORIGIN

  Query Match             65.9%;  Score 33.6;  DB 2;  Length 128529;
  Best Local Similarity   81.2%;  Pred. No. 3.8;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||  | |
Db       6736 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 6689


RESULT 12
AC008810
LOCUS       AC008810              157325 bp    DNA     linear   PRI 28-JUN-2003
DEFINITION  Homo sapiens chromosome 5 clone CTD-2096I23, complete sequence.
ACCESSION   AC008810
VERSION     AC008810.8  GI:32328942
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 157325)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2003) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jun 28, 2003 this sequence version replaced gi:14993661.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.3.
FEATURES             Location/Qualifiers
     source          1. .157325
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2096I23"
ORIGIN

  Query Match             65.9%;  Score 33.6;  DB 9;  Length 157325;
  Best Local Similarity   81.2%;  Pred. No. 3.6;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||  | |
Db     123706 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 123753


RESULT 13
AC093264
LOCUS       AC093264              164217 bp    DNA     linear   PRI 01-FEB-2002
DEFINITION  Homo sapiens chromosome 5 clone RP11-263G2, complete sequence.
ACCESSION   AC093264
VERSION     AC093264.2  GI:18464051
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 164217)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:15193398.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .164217
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-263G2"
ORIGIN

  Query Match             65.9%;  Score 33.6;  DB 9;  Length 164217;
  Best Local Similarity   81.2%;  Pred. No. 3.5;
  Matches   39;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 AAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCA 50
              | | | |||||||||||| || ||||| ||||||||||||||||  | |
Db      30833 AGATAAAAATTGGACTTCAAATTAAACACTTTTGTGCTTCAAAAGTTA 30880


RESULT 14
BS000239/c
LOCUS       BS000239              209016 bp    DNA     linear   PRI 07-OCT-2003
DEFINITION  Pan troglodytes chromosome 22 clone:PTB-153E07, map 22, complete
            sequences.
ACCESSION   BS000239
VERSION     BS000239.1  GI:37537506
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   The Chimpanzee Chromosome 22 Sequencing Consortium.
  TITLE     DNA sequence of chimpanzee chromosome 22 and its evolutionary
            implications
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 209016)
  AUTHORS   Tsai,S., Liu,T., Wu,K., Liao,T. and Hsiao,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2003) Shih-Feng Tsai, National Health Research
            Institutes (NHRI), Division of Molecular and Genomic Medicine; 128,
            Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan
            (E-mail:petsai@nhri.org.tw, URL:http://www.nhri.org.tw/,
            Tel:886-2-28267319, Fax:886-2-28200552)
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
```

```
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
            Center: National Yang Ming University Genome Research Center Center
            code: YMGC
            Web site: http://genome.ym.edu.tw/
            Contact: sequence@ym.edu.tw
            -------------- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: HI
            -------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly
            program: Phrap; version 0.990319
            Consensus quality:   207,750 bases at least Q40
            Consensus quality:   207,996 bases at least Q30
            Consensus quality:   208,014 bases at least Q20
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            --------------
            Source information:
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS145
            --------------
            Sequence Quality Assessment:
            This entry has been annotated with sequence
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than 1 error in
            10,000 bp.
            ----------------------
            Neighboring clones: PTB-152N20(left) and RP43-055A16(right).
FEATURES             Location/Qualifiers
     source          1. .209016
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="22"
                     /clone="PTB-153E07"
                     /clone_lib="PTB1 chimpanzee BAC"
     misc_feature    16459. .16463
                     /note="low quality region"
     misc_feature    43909. .44908
                     /note="gap containing unresolved di-nucleotide repeats,
                     (TG)n"
     misc_feature    45769. .45771
                     /note="low quality region"
     misc_feature    45776
                     /note="low quality region"
     misc_feature    45778
                     /note="low quality region"
     misc_feature    46232. .46235
                     /note="low quality region"
     misc_feature    46277. .46280
                     /note="low quality region"
     misc_feature    128655
                     /note="low quality region"
     misc_feature    128691
                     /note="low quality region"
ORIGIN

  Query Match             65.5%;  Score 33.4;  DB 9;  Length 209016;
  Best Local Similarity   78.4%;  Pred. No. 3.8;
  Matches   40;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACATCAT 51
              | | ||| ||||||| |||||| ||||| ||||||||||||  ||   |||
Db     123272 AAATAAATAAATTGGGCTTAAAATTAAAAACTTTTGTGCTTGCAAGGACAT 123222


RESULT 15
AL365334/c
LOCUS       AL365334               173053 bp    DNA     linear   ROD 29-JUN-2002
DEFINITION  Mouse DNA sequence from clone RP23-392F1 on chromosome 1, complete
            sequence.
ACCESSION   AL365334
VERSION     AL365334.13  GI:20068419
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 173053)
  AUTHORS   Blakey,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 7, 2002 this sequence version replaced gi:14141369.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-392F1 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            --------------.
FEATURES             Location/Qualifiers
     source          1. .173053
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="RP23-392F1"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             63.9%;  Score 32.6;  DB 10;  Length 173053;
  Best Local Similarity   80.9%;  Pred. No. 6.6;
  Matches   38;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy          1 ACAAAAAGAAATTGGACTTAAAGTTAAATACTTTTGTGCTTCAAACA 47
              ||||| || ||| | ||||| |||||||||||||| | | ||||||| |
Db     140706 ACAACAATAAAATCGACTTTAAGTTAAATACTTCTTTACTTCAAAAA 140660

Search completed: May  7, 2004, 14:30:58
Job time : 538.003 secs
```